IN PARTNERSHIP WITH:
Ecole nationale supérieure des techniques avancées

Activity Report 2012

Project-Team FLOWERS

Flowing Epigenetic Robots and Systems

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Robotics

Table of contents

Project-Team FLOWERS

Keywords: Developmental Robotics, Human-robot Interaction, Machine Learning, Language Acquisition, Robotics, Adaptive Systems, Sensorimotor Development

The Inria-ENSTA ParisTech FLOWERS project-team is bi-located between Inria in Bordeaux and ENSTA-ParisTech in Paris.

Creation of the Project-Team: July 01, 2008 *, Updated into Project-Team:* January 01, 2011 .

1. Members

Research Scientists
Pierre-Yves Oudeyer [Team Leader, Research Director (DR2), Inria, since january 2008, HdR]
Manuel Lopes [Research Associate (CR2), Inria, since october 2010]

Faculty Members
David Filliat [Professor, ENSTA-ParisTech, since june 2011, HdR]
Alexander Gepperth [Assistant Professor, ENSTA-ParisTech, since september 2011]
Freek Stulp [Assistant Professor, ENSTA-ParisTech, since november 2011]

Engineers
Jerome Bechu [Expert Engineer, since Oct. 2008, finished July 2012]
Haylee Fogg [Associate Engineer, since Oct. 2010, finished Oct. 2012]
Paul Fudal [Associate Engineer, since Oct. 2010]
Antoine Hoarau [Associate Engineer, since Oct. 2012]
Emmanuel Battesti [Associate Engineer, since Oct. 2012]

PhD Students
Pierre Rouanet [Inria, since Oct. 2008, finished March 2012]
Fabien Benureau [Université Bordeaux I, since Sept. 2010]
Thomas Cederborg [Inria, since Dec. 2009]
Jonathan Grizou [Inria, since Oct. 2011]
Matthieu Lapeyre [Inria, since Sept. 2010]
Olivier Mangin [Université Bordeaux I, since Sept. 2010]
Mai Nguyen [Inria, since Oct. 2010]
Natalia Lyubova [ENSTA-ParisTech, since november 2010]
Ievgen Perederieiev [Inria, since Oct. 2012]

Post-Doctoral Fellows
Thomas Degris [Inria, since Oct. 2011]
Clément Moulin-Frier [Inria, since Oct. 2011]

Administrative Assistant
Nicolas Jahier [Inria]

2. Overall Objectives

2.1. Introduction

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

These are the questions that are being investigated in the FLOWERS research team at Inria Bordeaux Sud-Ouest. Rather than trying to imitate the intelligence of adult humans like in the field of Artificial Intelligence, we believe that trying to reconstruct the processes of development of the child's mind will allow for more adaptive, more robust and more versatile machines. This approach is called developmental robotics, or epigenetic robotics, and imports concepts and theories from developmental psychology. As most of these theories are not formalized, this implies a crucial computational modeling activity, which in return provides means to assess the internal coherence of theories and sketch new hypothesis about the development of the human child's sensorimotor and cognitive abilities.

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

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

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

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

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

2.1.1. Research axis

The work of FLOWERS is organized around the following three axis:

- **Intrinsically motivated exploration and sensorimotor learning:** intrinsic motivation are mechanisms that have been identified by developmental psychologists to explain important forms of spontaneous exploration and curiosity. In FLOWERS, we try to develop computational intrinsic motivation systems, and test them on robots, allowing to regulate the growth of complexity in exploratory behaviours. These mechanisms are studied as active learning mechanisms, allowing to learn efficiently in large inhomogeneous sensorimotor spaces;

- **Cumulative learning of sensorimotor skills:** FLOWERS develops machine learning algorithms that can allow embodied machines to acquire cumulatively sensorimotor skills. In particular, we develop optimization and reinforcement learning systems which allow robots to discover and learn dictionaries of motor primitives, and then combine them to form higher-level sensorimotor skills.
- **Natural and intuitive social learning:** FLOWERS develops interaction frameworks and learning mechanisms allowing non-engineer humans to teach a robot naturally. This involves two sub-themes: 1) techniques allowing for natural and intuitive human-robot interaction, including simple ergonomic interfaces for establishing joint attention; 2) learning mechanisms that allow the robot to use the guidance hints provided by the human to teach new skills;
- **Discovering and abstracting the structure of sets of uninterpreted sensors and motors:** FLOWERS studies mechanisms that allow a robot to infer structural information out of sets of sensorimotor channels whose semantics is unknown, for example the topology of the body and the sensorimotor contingencies (propriocetive, visual and acoustic). This process is meant to be open-ended, progressing in continuous operation from initially simple representations to abstract concepts and categories similar to those used by humans.

2.2. Highlights of the Year

2.2.1. Ergo-Robots: Large-scale life-long learning robot experiment

The FLOWERS team, in collaboration with University Bordeaux I/Labri, has participated as a central actor of the exhibition “Mathematics: A Beautiful Elsewhere” at Fondation Cartier pour l’Art Contemporain in Paris. This installation, called “Ergo-Robots/FLOWERS Fields” was made in collaboration with artist David Lynch and mathematician Mikhail Gromov (IHES, France), and shows computational models of curiosity-driven learning, human-robot interaction as well as self-organization of linguistic conventions. This exhibition, at the crossroads of science and art, allowed to disseminate our work towards the general public, explaining concepts related to learning mechanims in humans and robots to a large audience (80000 visitors). This was also an opportunity for experimenting and improving our technologies for life-long robot learning experimentation. For one of the first times in the world outside the laboratory, we demonstrated how it is possible to achieve experimentation with learning robots quasi-continuously for 5 months. This opens novel stimulating scientific perspectives in the field of developmental robotics. This experimentation was presented through large audience radios, magazines and newspapers (France Inter, France Culture, RFI, Sciences et Avenir, Tangente, Financial Times, Daily Telegraph, Libération, ...).
More information available at: http://flowers.inria.fr/ergo-robots.php and http://fondation.cartier.com/.

2.2.2. MACSi: Integrated system for curiosity-driven visual object discovery on ICub robot

In the frame of the MACSi ANR project conducted together with ISIR (UPMC - Paris) a complete cognitive architecture for humanoid robots interacting with objects and caregivers in a developmental robotics scenario has been integrated on the iCub robot [43]. The architecture is foundational to the MACSi project and to several research axis of FLOWERS: it is designed to support experiments to make a humanoid robot gradually enlarge its repertoire of known objects and skills combining autonomous learning, social guidance and intrinsic motivation. This complex learning process requires the capability to learn affordances, i.e. the capacity for the robot to predict which actions are possible on scene elements. Several papers presenting the general framework for achieving these goals, focusing on the elementary action, perception and interaction modules have been published. This architecture is an important milestone of the project, enabling future experiments on object learning and recognition, object categorization and interaction between autonomous exploration and social guidance.

2.2.3. Algorithmic architecture for learning inverse models in high-dimensional robots

Through the design of the SAGG-RIAC algorithmic architecture, and its publication in a major robotics journal [22], we have produced a highly-efficient system for intrinsically motivated goal exploration mechanism which allows active learning of inverse models in high-dimensional redundant robots. Based on active goal babbling, this allows a robot to efficiently and actively learn distributions of parameterized motor skills/policies that solve a corresponding distribution of parameterized tasks/goals. We have conducted experiments with high-dimensional continuous sensorimotor spaces in three different robotic setups: 1) learning the inverse kinematics in a highly-redundant robotic arm, 2) learning omnidirectional locomotion with motor primitives in a quadruped robot, 3) an arm learning to control a fishing rod with a flexible wire. We showed that 1) exploration in the task space can be a lot faster than exploration in the actuator space for learning inverse models in redundant robots; 2) selecting goals maximizing competence progress creates developmental trajectories driving the robot to progressively focus on tasks of increasing complexity and is statistically significantly more efficient than selecting tasks randomly, as well as more efficient than different standard active motor babbling methods; 3) this architecture allows the robot to actively discover which parts of its task space it can learn to reach and which part it cannot.

2.2.4. Formalization of several links between intrinsic motivation architectures and statistical machine learning

We incorporated several key concepts of intrinsically motivated developmental learning, especially measures of learning progress for curiosity-driven exploration, in several standard machine learning formalisms. First, we introduced and formalized a general class of learning problems for which a developmental learning strategy is optimal [47]. This class of learning problems characterizes problems where the issue of life-long multitask learning under bouded ressources is crucial. Within this formalization, we related the SAGG-RIAC architecture [22] with multi-armed bandits formalisms [47] allowing to study the properties of problems where there several discrete choices to make to accelerate learning. Third, we also included empirical measures of learning progress in standard reinforcement learning problem allowing to automatically choose the best exploration strategy [42] and to extend Rmax approaches, for exploration in model-based RL, to non-stationary problems [46].

2.2.5. Bridging black-box optimization and RL for skill learning in robots

In this year, we have made substantial advances in understanding of the relationship between black-box optimization and reinforcement learning for direct policy search, and the application of such methods to robotics manipulation, as well as their use for modelling human behavior. The key discovery has been that black-box optimization and reinforcement learning have converged towards a same set of algorithmic properties, such as parameter perturbation and reward-weighted averaging, allowing for a direct comparison and integration of such algorithms (see "Relationship between Black-Box Optimization and Reinforcement Learning" below). On the one hand, this has enabled us to exploit principles from black-box optimization, such as covariance matric adaptation, in the context of reinforcement learning. The resulting algorithm (PI^2-CMAES) enables adaptive exploration and life-long learning in robots [63], and in reaching experiments leads to proximo-distal maturation as an emergent property [60] (see "Emergent Proximo-Distal Maturation through Adaptive Exploration" below). On the other hand, it has allowed us to demonstrate that black-box optimization outperforms reinforcement learning for a particular class of policies [69]. This is an important result, as these types of policies are typically used for robotic skill learning. Therefore, more efficient and robust black-box optimization algorithms may be applied to learning with such policies, *without* compromising convergence speed and cost of the final solution.

2.2.6. Algorithm for learning sequences of motion primitives

As for applications, we have also extended policy improvement algorithms to work with sequences of motion primitives, enabling 11-DOF manipulation robots to learn how to grasp under uncertainty through fine manipulation, and perform extended pick-and-place tasks [31] (see "Reinforcement Learning with Sequences of Motion Primitives for Robust Manipulation" below). We have also shown that learning variable impedance control is able to mimic the behavior of humans when exposed to force fields (see “Model-free Reinforcement Learning of Impedance Control in Stochastic Environments” below) .

2.2.7. Algorithms for autonomous dimensionality reduction

In 2012, we have made significant progress in incremental online learning algorithms capable of finding latent variables in high-dimensional sensory spaces, by either using the principle of multimodal correspondence[24] or weak, self-generated supervision[40]. These advances will be key in further extending the applicability of key artificial curiosity algorithms for learning with high-dimensional sensori spaces.

The following paper obtained the Best Paper Award in the category “Computational Models of Cognitive Development” at the IEEE ICDL-Epirob international conference: [53]
BEST PAPER AWARD :
[53] **Curiosity-driven phonetic learning in ICDL-Epirob - International Conference on Development and Learning, Epirob**. C. MOULIN-FRIER, P.-Y. OUDEYER.

3. Scientific Foundations

3.1. Scientific Foundations

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 grand challenge is thus to be able to build robotic 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" [120]. 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 [90], 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 [97] [122]. The approach of developmental robotics consists in importing and implementing concepts and mechanisms from developmental psychology [102], cognitive linguistics [79], and developmental cognitive neuroscience [93] 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 [76] [110], grounding [88], situatedness [70], self-organization [118] [111], enaction [121], and incremental learning [77].

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 [14]. Thus, exploration needs to be organized and guided. The approach of epigenetic robotics is to take inspiration from the mechanisms that allow human infants to be progressively guided, i.e. to develop. There are two broad classes of guiding mechanisms which control exploration:

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

3.1.1. Internal guiding mechanisms

In infant development, one observes a progressive increase of the complexity of activities with an associated progressive increase of capabilities [102], 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" [72] [84]. 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 [74] [80] [82]. 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 behaviors and curiosity [81] [91] [116]. Based on this, a number of researchers have began in the past few years to build computational implementation of intrinsic motivation [14] [108] [114] [73] [92] [100] [115]. 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 [108][14] [109] [113]. 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 [78] [89] typically lead to catastrophic results. The issue of high dimensionality does not only concern motor spaces, but also sensory spaces, leading to the problem of correctly identifying, among typically thousands of quantities, those latent variables that have links to behavioral choices. In FLOWERS, we aim at developing intrinsically motivated exploration mechanisms that scale in those spaces, by studying suitable abstraction processes in conjunction with exploration strategies.

3.1.2. Socially Guided and Interactive Learning

Social guidance is as important as intrinsic motivation in the cognitive development of human babies [102]. 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 [71]. 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 [119], [75], motivated by the various mechanisms that allow humans to socially guide a robot [112]. In an interactive context the steps of self-exploration and social guidances are not separated and a robot learns by self exploration and by receiving extra feedback from the social context [119], [94] [101].

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 [104], [106]. We also work on developing perceptual capabilities by gradually segmenting the perceptual space and identifying objects and their characteristics through interaction with the user [48] and robots experiments [43]. 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 [96].

4. Application Domains

4.1. Application Domains

- **Personal robotics.** Many indicators show that the arrival of personal robots in homes and everyday life will be a major fact of the 21st century. These robots will range from purely entertainment or educative applications to social companions that many argue will be of crucial help in our aging society. For example, UNECE evaluates that the industry of entertainment, personal and service robotics will grow from $5.4Bn to $17.1Bn over 2008-2010. Yet, to realize this vision, important obstacles need to be overcome: these robots will have to evolve in unpredictable homes and learn new skills while interacting with non-engineer humans after they left factories, which is out of reach of current technology. In this context, the refoundation of intelligent systems that developmental robotics is exploring opens potentially novel horizons to solve these problems.

- **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.
- **Video games.** In conjunction with entertainment robotics, a new kind of video games are developing in which the player must either take care of a digital creature (e.g. Neopets), or tame it (e.g. Nintendogs), or raise/accompany them (e.g. Sims). The challenges entailed by programming these creatures share many features with programming personal/entertainment robots. Hence, the video game industry is also a natural field of application for FLOWERS.
- **Environment perception in intelligent vehicles.** When working in simulated traffic environments, elements of FLOWERS research can be applied to the autonomous acquisition of increasingly abstract representations of both traffic objects and traffic scenes. In particular, the object classes of vehicles and pedestrians are if interest when considering detection tasks in safety systems, as well as scene categories (”scene context”) that have a strong impact on the occurrence of these object classes. As already indicated by several investigations in the field, results from present-day simulation technology can be transferred to the real world with little impact on performance. Therefore, applications of FLOWERS research that is suitably verified by real-world benchmarks has direct applicability in safety-system products for intelligent vehicles.
- **Automated Tutoring Systems.** Optimal teaching and efficient teaching/learning environments can be applied to aid teaching in schools aiming both at increase the achievement levels and the reduce time needed. From a practical perspective, improved models could be saving millions of hours of students' time (and effort) in learning. These models should also predict the achievement levels of students in order to influence teaching practices.

5. Software

5.1. Perception Tools

Participants: David Filliat [correspondant], Natalia Lyubova, Louis-Charles Caron, Alexander Gepperth.

5.1.1. Perception Abstraction Engine

Participants: David Filliat [correspondant], Natalia Lyubova.

PAE (Perception Abstraction Engine) is a C++ library developed to provide a uniform interface to existing visual feature detector such as SIFT, SURF, MSER, superpixels, etc... Its main goal is to be able to use these various feature detectors in a "bag of feature" approach for applications such as robot localisation and object recognition. Several approach are also implemented for the visual vocabularies, in particular the fast incremental vocabularies developed in the team.

The library provide common C++ interfaces to feature detectors, visual features and visual vocabularies. A factory approach make it possible to change the feature detectors and visual vocabularies types and parameters through configuration strings, without the need to recompile. Some applications are also included in the library, in particular topological robot localization (room recognition) and visual object recognition. An Urbi interface is also provided for these modules.

5.1.2. Incremental object discovery

Participants: Natalia Lyubova [correspondant], David Filliat.

This software makes it possible to detect, model and recognize objects in a scenario of interaction between a humanoid robot and a human teacher. It is based either on standard images, or on the kinect camera to take advantage of the depth information. The software is written in C++ and relies mainly on PAE and OpenCV.

The software implements several modules: candidate object segmentation based on motion information, keypoint-based object tracking, incremental object model construction integrating multiple features (keypoints + superpixels) and object categorisation based on mutual information with robot motors (making it possible to segment robot parts, objects and humans).

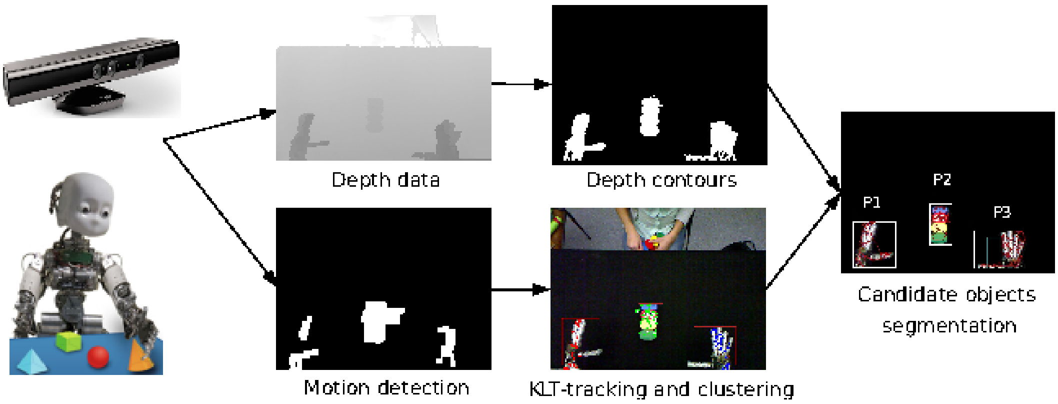


Figure 1. System Overview of the Incremental object discovery Software.

5.1.3. Object recognition from a 3-D point cloud

Participants: Louis-Charles Caron [correspondant], Alexander Gepperth, David Filliat.

This software scans the 3-D point cloud of a scene to find objects and match them against a database of known objects. The process consists in 3 stages. The segmentation step finds the objects in the point cloud, the feature extraction computes discriminating properties to be used in the classification stage for object recognition.

The segmentation is based on simple assumptions about the geometry of an indoor scene. Successive RANSACs are used to find large planes, which correspond to the floor, ceiling and walls. The cloud is stripped from the points belonging to these planes. The remaining points are clustered, meaning that close-by points are considered to form a single object.

Objects are characterized by their shape and color. Color histograms and SIFT features are computed, using the PAE library, to capture the visual appearance of the objects. Their shape is encoded by computing thousands of randomly chosen SURFLET features to construct a relative frequency histrogram.

An early classification is done using each of the 3 features separately. For the color features a bag of words approach (from PAE) is used. For the shape feature, the minimum squared distance between the object's histogram and that of all objects in the database is calculated. Classification scores are then fused by a feed-forward neural network to get the final result [39].

5.1.4. PEDDETECT: GPU-accelerated person detection demo

Participant: Alexander Gepperth [correspondant].

PEDDETECT implements real-time person detection in indoor or outdoor environments. It can grab image data directly from one or several USB cameras, as well as from pre-recorded video streams. It detects mulitple persons in 800x600 color images at frame rates of >15Hz, depending on available GPU power. In addition, it also classifies the pose of detected persons in one of the four categories "seen from the front", "seen from the back", "facing left" and "facing right". The software makes use of advanced feature computation and nonlinear SVM techniques which are accelerated using the CUDA interface to GPU programming to achieve high frame rates. It was developed in the context of an ongoing collaboration with Honda Research Institute USA, Inc.

5.2. Datasets

5.2.1. Choreography dataset

Participants: Olivier Mangin [correspondant], Haylee Fogg.

This database contains choreography motions recorded through a kinect device. These motions have a combinatorial structure: from a given set of primitive dance motions, choreographies are constructed as simultaneous execution of some of these primitive motions. Primitive dance motions are chosen from a total set of 48 motions and are spanned over one or two limbs, either the legs (e.g. walk, squat), left or right arm (e.g. wave hand, punch) or both arms (e.g. clap in hands, paddle). Complex choreographies are produced as the simultaneous demonstration of two or three of these primitive motion: either one for legs and one for both arm, or one for legs and one for each arm. The dataset has been used in the experiments from [52] for studying learning techniques allowing to identify dictionaries of motion primitives, and is publicly available at https://flowers.inria.fr/choreography_database.html.

5.3. Learning algorithms

5.3.1. RLPark - Reinforcement Learning Algorithms in JAVA

Participant: Thomas Degris [correspondant].

RLPark is a reinforcement learning framework in Java. RLPark includes learning algorithms, state representations, reinforcement learning architectures, standard benchmark problems, communication interfaces for three robots, a framework for running experiments on clusters, and real-time visualization using Zephyr. More precisely, RLPark includes:

- Online Learning Algorithms: Sarsa, Expected Sarsa, Q-Learning, On-policy and off-policy Actor-Critic with normal distribution (continuous actions) and Boltzmann distribution (discrete action), average reward actor-critic, TD, TD(λ), GTD(λ), GQ(λ), TDC
- State Representations: tile coding (with no hashing, hashing and hashing with mumur2), Linear Threshold Unit, observation history, feature normalization, radial basis functions
- Interface with Robots: the Critterbot, iRobot Create, Nao, Puppy, Dynamixel motors
- Benchmark Problems: mountain car, swing-up pendulum, random walk, continuous grid world

An example of RLpark running an online learning experiment on a reinforcement learning benchmark problem is shown in Figure 2.

RLPark was started in spring 2009 in the RLAI group at the university of Alberta (Canada) when Thomas Degris was a postdoc in this group. RLPark is still actively used by RLAI. Collaborators and users include Adam White, Joseph Modayil and Patrick Pilarski (testing) from the University of Alberta.

RLPark has been used by Richard Sutton, a professor and iCORE chair in the department of computing science at the University of Alberta, for a demo in his invited talk *Learning About Sensorimotor Data* at the Neural Information Processing Systems (NIPS) 2011 [1]. Patrick Pilarski used RLPark for live demos on television (Breakfast Television Edmonton, CityTV, June 5th, 2012) and at TEDx Edmonton on Intelligent Artificial Limbs[2]. So far, RLPark has been used in more than a dozens of publications (see http://rlpark.github.com/publications.html for a list).

[1] http://webdocs.cs.ualberta.ca/~sutton/Talks/Talks.html#sensorimotor
[2] http://www.youtube.com/watch?v=YPc-Ae7zqSo

RLPark has been ported to C++ by Saminda Abeyruwan, a student of the University of Miami (United States of America). The Horde architecture in RLPark has been optimized for GPU by Clément Gehring, a student of the McGill University in Montreal (Canada).

Future developments include the implementation of additional algorithms (the Dyna architecture, back propagation in neural networks, ...). A paper is under review for the JMLR Machine Learning Open Source Software. Documentation and tutorials are included on the RLPark web site [3]. RLPark is licensed under the open source Eclipse Public License.

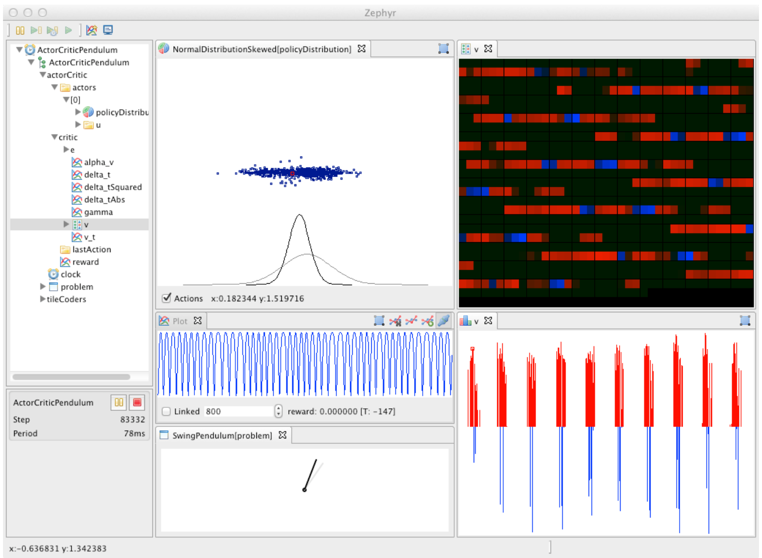


Figure 2. An example of an experiment in RLPark. Zephyr displays two views of a learned weight vector, an animation of the problem, the current policy distribution learned by the algorithm and the reward obtained by the algorithm. Videos are available at: http://rlpark.github.com.

5.3.2. DMP-BBO Matlab library

Participant: Freek Stulp [correspondant].

The dmp_bbo (Black-Box Optimization for Dynamic Movement Primitives) Matlab library is a direct consequence of the insight that black-box optimization outperforms reinforcement learning when using policies represented as Dynamic Movement Primitives. It implements several variants of the PI^{BB} algorithm for direct policy search. It is currently being used and extended by several FLOWERS members (Manuel Lopes, Clement Moulin-Frier) and external collaborators (Jonas Buchli, Hwangbo Jemin of ETH Zurich). This code was used for the following publications: [63], [60], [62].

5.3.3. PROPRE: simulation of developmental concept formation using PYTHON

Participant: Alexander Gepperth [correspondant].

This simulation software implements the algorithms described in [24], [40]. It is available online under the URL www.gepperth.net/downloads.html. The simulation is implemented in PYTHON for easy use, yet the time-critical core functions are written in C.

[3] http://rlpark.github.com

5.3.4. pySteamPlayer: synchronized replay of multiple sensor recordings and supplementary data

Participant: Alexander Gepperth [correspondant].

This Python software is intended to facilitate the application of machine learning algorithms by avoiding to work directly with an embodied agent but instead with data recorded in such an agent. Assuming that non-synchronous data from multiple sensors (e.g., camera, Kinect, laser etc.) have been recorded according to a flexible format defined by the pyStreamPlayer architecture, pyStreamPlayer can replay these data while retaining the exact temporal relations between different sensor measurements. As long as the current task does not involve the generation of actions, this software allows to process sensor data as if it was coming from an agent which is usually considerably easier. At the same time, pyStreamPlayer allows to replay arbitrary supplementary information such as, e.g., object information, as if it was coming from a sensor. In this way, supervision information can be stored and accessed together with sensory measurements using an unified interface. pyStreamPlayer has been used to facilitate real-world object recognition tasks, and several of the major databases in this field (CalTech Pedestrian database, HRI RoadTraffic traffic objects database, CVC person database, KITTI traffic objects database) have been converted to the pyStreamPlaer format and now serve as a source of training and test data for learning algorithms.

pyStreamPlayer has been integrated into a ROS node as well, allowing th replay and transmission across networks of distributed processes.

5.4. Software Platforms

5.4.1. Robust robotics manipulation - Object detection and tracking

Participants: Antoine Hoarau [ADT Engineer Since Nov. 2012], Freek Stulp [Supervisor], David Filliat [Supervisor].

Autonomous human-centered robots, for instance robots that assist people with disabilities, must be able to physically manipulate their environment. There is therefore a strong interest within the FLOWERS team to apply the developmental approach to robotics in particular to the acquisition of sophisticated skills for manipulation and perception. ENSTA-ParisTech has recently acquired a Meka (cf. 3) humanoid robot dedicated to human-robot interaction, and which is perfectly fitted to this research. The goal of this project is to install state-of-the-art software architecture and libraries for perception and control on the Meka robot, so that this robot can be jointly used by FLOWERS and ENSTA. In particular, we want to provide the robot with an initial set of manipulation skills.
The goal is to develop a set of demos, which demonstrate the capabilities of the Meka, and provide a basis on which researchers can base their experiments. As the robot is not yet available at ENSTA, initial work focused on the robot's environment, meaning ROS and the M3 software (provided by Meka Robotics, based on both C++ and Python scripts) and on trying to implement a simple ball-catching demo : the idea is to throw a ball toward the robot which catch it (basic human-robot interaction, combining both perception and control). Different tracking algorithms are being tried for the ball, such as Camshift, Hough Circles + Kalman Filter, or more complex LineMod (all included in OpenCV) to finally estimate its trajectory for the robot to catch it. The M3 software provided by Meka Robotics contains a simulation environment that allows us to work without the robot hardware (cf. 4.

5.4.2. ErgoRobot/Flowers Field Software

Participants: Jérôme Béchu [correspondant], Pierre-Yves Oudeyer, Pierre Rouanet, Olivier Mangin, Fabien Benureau, Mathieu Lapeyre.

Figure 3. The Meka robot plateform acquired by ENSTA ParisTech

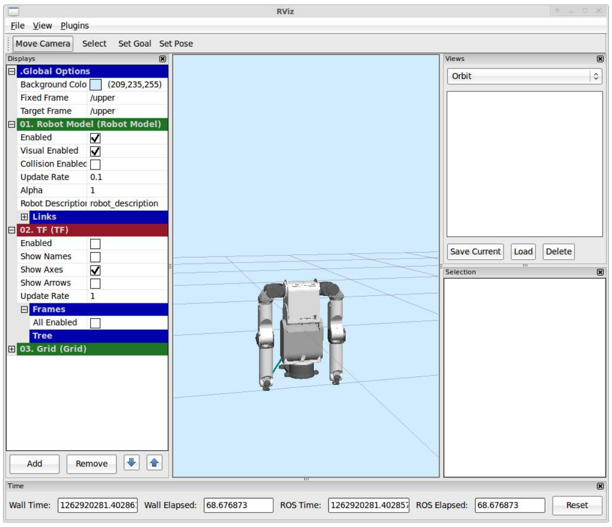


Figure 4. M3 simulation through ROS and Rviz

In the context of its participation to the exhibition “Mathematics: A Beautiful Elsewhere” at Fondation Cartier pour l’Art Contemporain in Paris (19th October 2011 to 18th March 2012), the team has elaborated and experimented a robotic experimental set-up called “Ergo-Robots/FLOWERS Fields”. This set-up is not only a way to share our scientific research on curiosity-driven learning, human-robot interaction and language acquisition with the general public, but, as described in the Results and Highlights section, attacks a very important technological challenge impacting the science of developmental robotics: How to design a robot learning experiment that can run continuously and autonomously for several months?

The global scenario for the robots in the installation/experiment is the following. In a big egg that has just opened, a tribe of young robotic creatures evolves and explores its environment, wreathed by a large zero that symbolizes the origin. Beyond their innate capabilities, they are outfitted with mechanisms that allow them to learn new skills and invent their own language. Endowed with artificial curiosity, they explore objects around them, as well as the effect their vocalizations produce on humans. Human, also curious to see what these creatures can do, react with their own gestures, creating a loop of interaction which progressively self-organizes into a new communication system established between man and ergo-robots.

We now outline the main elements of the software architectures underlying this experimental setup.

5.4.2.1. System components

The software architecture is organized to control the experiment at several levels, and in particular:

- **Scenes**: The organization of behavioural scenes, managing the behaviours that are allowed to each robot at particular times and in particular contexts;
- **Behaviours**: The individual behaviours of robots, also called stems, which are outlined in the next section;
- **stems**: The low-level actions and perceptin of robots while executing their behaviours, including motors control on the five physical stems, color and intensity of lights inside the stem head, production of sounds through speakers. Sensors are the kinect used to interact with visitors, and motor feedback capabilities.

In addition to that a video projector is used to display some artistic view of stem agents internal state.

Figure 5. Three importants concepts in ErgoRobots

5.4.2.2. Behaviours

A number of innate behaviours were designed and are used by the robots as elementary behaviours of more complex behaviours, including the three following learning behaviours.

The Naming Game is a behaviour played by stems two-by-two and based on computational models of how communities of language users can self-organize shared lexicons. In the naming game, stems interact with each other in a stylised interaction. Repeated interactions lead to the development of a common repertoire of words for naming objects. More precisely, object belong to meaning spaces. Two such spaces have been implemented for the exhibition. The first one is related to object spatial categorization and the second one is related to movement categorization. The object space contains stems, some hole in walls and the interaction zone. The movement space contains representations of small dances that stem can produce and reproduce.

Object Curiosity is a behaviour in controlling intrinsically motivated exploration of the physical environnement by the stems. A small wood object is present in the reachable physical environement of the stem, attached on the top of a spring so that it is guaranteed that it comes back to its original position. The stem uses a motor primitive to act on the object and motor feedback to detect movements of the object. The robot learns through active exploration what kind of parameters motor primitive will result in touching the object.

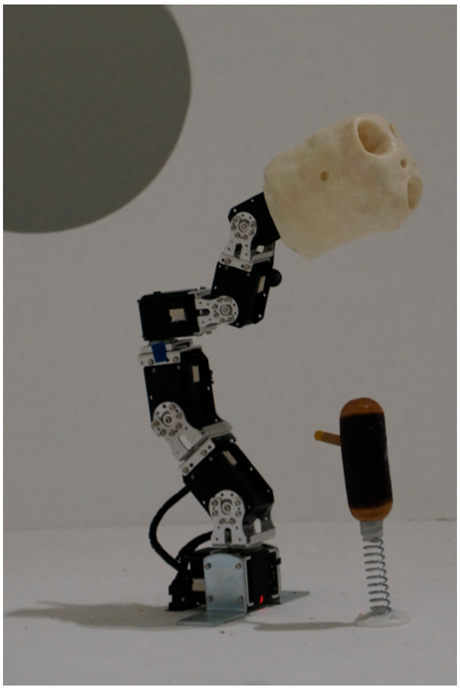


Figure 6. A Stem with the head designed by David Lynch and an Object

Birds Curiosity is a behaviour that drives robots to explore, through curiosity-driven learning, interaction with humans. One stem, generally the stem in the center, plays a sound, predicts the visitor reaction, look the interaction zone and wait the gesture of the visitor. To produce a sound the visitor have to make a gesture in space. In the next iterations, the robot chooses to produce sounds to human which produce most surprising responses from the human (i.e. the robot is “interested” to explore sound interactions which are not easily predictable by itself).. As describe in the picture, the space is split in four. Each zone corresponding with a sound.

5.4.2.3. Programming tools

The system is based on URBI and used some UObjects from UFlow. The most important part of the system is written in URBI script. Python and freenect [4] are used too.

The system at the startup detects motors and lights. It create dynamically a list of Stem. A Stem is one robot with 6 motors as described in hardware part.

[4]Kinect library

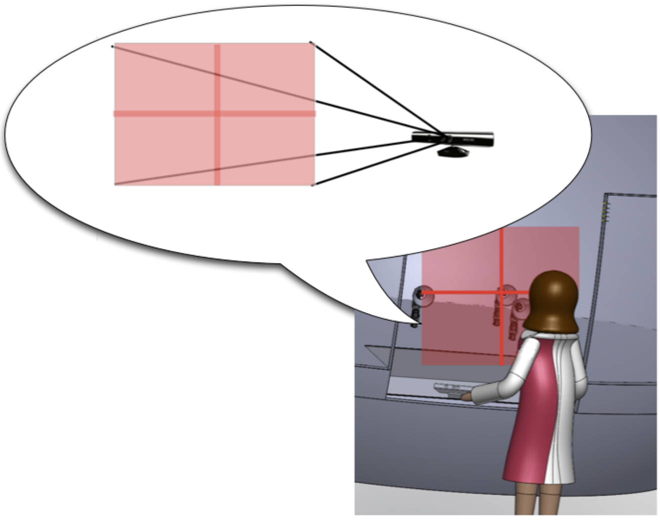


Figure 7. A virtual visitor interact with a virtual grid

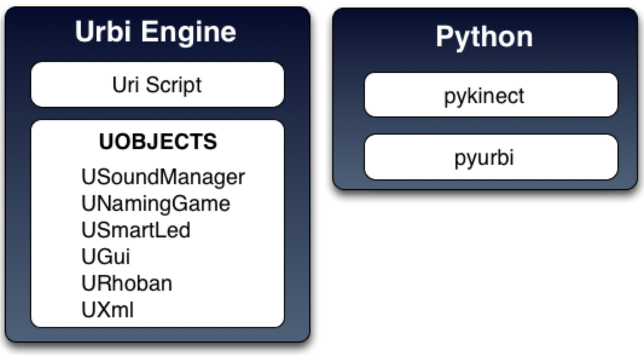


Figure 8. List of software used in ErgoRobots

To interact with people, we used the freenect library to interface with the kinect, with a binding to python where detection and following of gestures is made.

For the display, we display an abstract rendering of the structure inside each ErgoRobot, using a python parser to read and parse log file from the ErgoRobot system, and the Bloom/Processing software to create and display the rendering. Currently, the system has three displays, one for the naming game, another one for birds curiosity and the last one for objects curiosity.

The sound system used the UObject USoundManager. It plays sounds when required by a behaviour, it also plays word sounds in Naming Game behaviour.

The Light system used Linkm technologies. In the head of each ErgoRobot we put two lights devices. Each light device is a RGB Light. We can control the intensity of each primary color through I2C control. To control lights we used LinkM USB Device. And finally we used an UObject dedicated to communicate with the USB Device.

5.4.2.4. Maintenance

A dedicate maintenance software is used to switch off, switch on the system. This software is written in Python (and Qt). The status of ErgoRobots is display on the graphical interface. Buttons are present too : Start, Stop, Reset and Take a video.

Recently we added a video system to have a visual feedback of motors usage and also to detect eventual problems. This is a screenshot of the application :

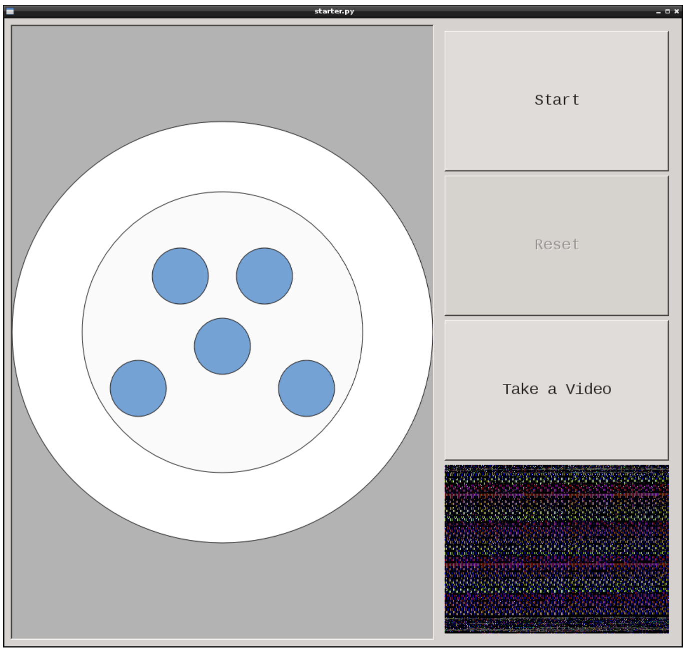


Figure 9. Maintenance Software for the ErgoRobots.

5.4.3. MonitorBoard - Complete solution for monitoring Rhoban Project robots

Participants: Paul Fudal [correspondant], Olivier Ly, Hugo Gimbert.

In collaboration with Rhoban Project/LaBRI/CNRS/Univ. Bordeaux I, the Flowers team took part in a project to exhibit robots at the International Exhibition in Yeosu - 2012 - South Korea (8 millions of visitors expected, from more than 100 countries). The installation consisted in three humanoids (one dancing, two playing on a spring) and five musicians (arms only) playing musical instruments (electric guitar, electric bass guitar, keytar, drums, DJ turntables). In order to increase the robustness of the robotic platform, a complete solution of software and hardware was build. The software solution aims to allow all robots to run safely during the whole exhibition (12 hours per days) and to provide an easy way to diagnose and identify potential electronic and mechanical failures. This software is able to monitor all robots at the same time, verify the health of each motors and each embedded systems. It is able to shutdown or reboot a robot if necessary using PowerSwitches (electric plugs controlled over network) and notify maintenance personal by email explaining the failure. All information is also logged for statistical use. This solution allows to monitor the whole platform without being present, and provides warning signs enabling preventive actions to be taken before an actual failures. It was entirely written in $C\#$ using Microsoft Visual Studio 2010 with .NET API and combined with the existing Rhoban Project API, extended and modified for this purpose. It also involved electric plugs controlled over a network connection.

5.4.4. Motor tracking system

Participants: Jérôme Béchu, Olivier Mangin [correspondant].

We developed a website interface to a database of motors used to build robots in the team. This system is designed for internal use in the team and was developed using the django web framework (https://www.djangoproject.com/) .

5.5. Visualization Tools

5.5.1. Zephyr - Realtime Visualization in JAVA

Participant: Thomas Degris [correspondant].

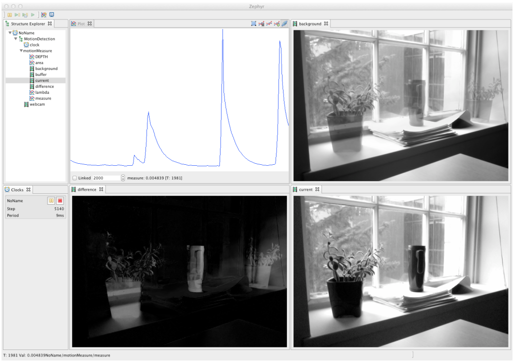
Zephyr is a software to visualize numeric variables and data structure in real time and at different time scale. Zephyr is practical because it requires only minimal changes in the code: it uses Java reflexivity to automatically detect variables in the code to monitor and data structure with an associated dedicated view. Zephyr can easily be extended with new plugins because it is based on the popular Eclipse Rich Client Platform. Consequently, Zephyr takes advantage of an already existing and fully operational Eclipse plugins for many of its functionalities. Finally, Zephyr is distributed with a Java python virtual machine named Jython and a lisp implementation named Clojure. An example of a Zephyr screen is shown in Figure 10.

Zephyr was started in fall 2009 in the RLAI group at the university of Alberta (Canada) when Thomas Degris was a postdoc in this group. Zephyr is still actively used by RLAI. Users include Adam White, Joseph Modayil and Patrick Pilarski from the University of Alberta. Zephyr has been registered on the Eclipse marketplace since October 2011. Documentation about Zephyr is included on its website: http://zephyrplugins.github.com. Zephyr is licensed under the open source Eclipse Public License.

5.6. Hardware

5.6.1. Poppy Platform

Participant: Matthieu Lapeyre [correspondant].



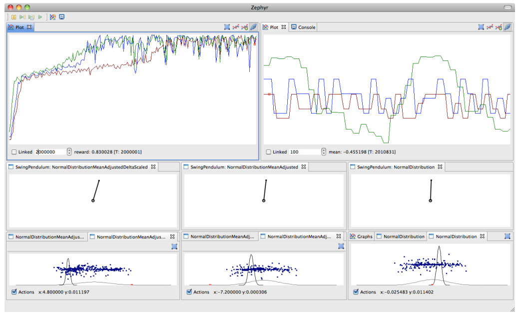


Figure 10. Left: Zephyr showing the different steps of a video processing pipeline in real-time. Right: Zephyr showing different data structure and variables of a reinforcement learning agent at different time scale. A video is available at: http://zephyrplugins.github.com.

5.6.1.1. Main goals :

No current platform (Nao [86], Darwin Op [87], Nimbro Op [117], HRP-2, ...) does offer both a adapted morphology in the sense of allowing physical interaction (safe, compliant, playful) and optimized for walking. So to explore these challenges we have decided to build a new bio-inspired humanoid robotic platform, called Poppy, which provides some of the software and hardware features needed to explore both social interaction and biped locomotion for personal robot. It presents the following main features to make it an interesting platform to study how the combination of morphology and social interaction can help the learning:

- Design inspired from the study of the anatomy of the human body and its bio-mechanic
- Dynamic and reactive: we try to keep the weight of the robot as low as possible (geometry of the pieces and smaller motors)
- Social interaction: screen for communication and permits physical interaction thanks to compliance
- Study of the morphology of the leg to improve the biped walking
- Practical platform: low cost, ease of use and easy to reproduce

5.6.1.2. Overview :

Poppy platform (Figure 11) is a humanoid, it is 84cm tall for 3 kg. It has a large sensor motors space including 25 dynamical motors (MX-28 and AX-12), force sensors under its feet and some extra sensors in the head: 2 HD-wide angle-cameras, stereo-micros and an inertial central unit (IMU 9DoF) plus a large LCD Screen (4 inch) for visual communication (e.g. emotions, instructions or debug). The mechanical parts were designed and optimized to be as light as possible while maintaining the necessary strength. For this, the choice of a lattice beam structure manufactured with 3Dprinting polyamide was used.

The poppy morphology is designed based on the actual human body. We have deeply studied the biomechanics of the human body and have extracted some interesting features for humanoid robotics. This inspiration is expressed in the whole structure (e.g. the limb proportions) and in particular in the trunk and legs.

Poppy uses the bio-inspired trunk system introduced by Acroban [98]. These five motors allow it to reproduce the main changes brought by the human spine. This feature allows the integration of more natural and fluid motion while improving the user experience during physical interactions. In addition, the spine plays a fundamental role in bipedal walking and postural balance by actively participating in the balancing of the robot.

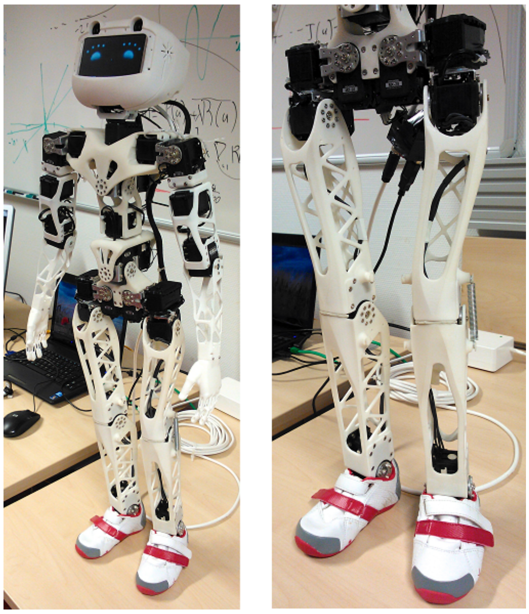


Figure 11. a. Global view of the Poppy platform. b. Zoom on legs design

The legs were designed to increase the stability and agility of the robot during the biped walking by combining bio-inspired, semi-passive, lightweight and mechanical-computation features. We will now describe two examples of this approach:

The architecture of the hips and thighs of Poppy uses biomechanical principles existing in humans. The human femur is actually slightly bent at an angle of about 6 degrees. In addition, the implantation of the femoral head in the hip is on the side. This results in a reduction of the lateral hip movement needed to move the center of gravity from one foot to another and a decrease in the lateral falling speed. In the case of Poppy, the inclination of its thighs by an angle of 6 degrees causes a gain of performance of more than 30% for the two above mentioned points.

Another example is Poppy's feet. Poppy has the particularity of having small feet compared to standard humanoids. It has humanly proportioned feet (ie about 15% of its total size).It is also equipped with compliant toes joints (see Figure 12.a). We believe that this feet involve two keys features to obtain a human-like and efficient walking gait. However, that raises problems regarding balance because the support polygon is reduced. We decided to add pressure sensors under each foot in order to get accurate feedback of the current state of the robot (see Figure 12.b).

5.6.1.3. Future works :

In our work, we explore the combination of both a bio-inspired body and bio-inspired learning algorithms. We are currently working on experiments involving Poppy to perform skill learning. First we would like to succeed in achieving an effective postural balance using the articulated spine, the feet pressure sensors and the IMU. Then, we would like to perform experiments on the learning of biped walking using algorithms such as the ones described in [95] or [83]. We are expecting to clearly reduce the learning time needed and increase the quality of the learned tasks thanks to the bio-inspired morphology of Poppy.

We are also interested in social interactions with non-expert users. We would like to conduct user study to evaluate how playful physical interactions and emotions could improve learning in robotics. We think that the poppy platform could be very suitable for such studies.

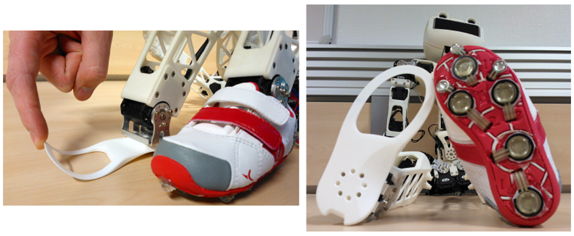


Figure 12. Poppy feet use actual children shoes combine with a compliant feet, toes (a.) and pressure sensors (b.)

5.6.2. Ergo-Robots/FLOWERS Fields: Towards Large-Scale Robot Learning Experiments in the Real World

Participants: Jerome Bechu, Fabien Benureau, Haylee Fogg, Paul Fudal, Hugo Gimbert, Matthieu Lapeyre, Olivier Ly, Olivier Mangin, Pierre Rouanet, Pierre-Yves Oudeyer.

In the context of its participation to the exhibition “Mathematics: A Beautiful Elsewhere” at Fondation Cartier pour l’Art Contemporain in Paris, starting from 19th October 2011 and to be held until 18th March 2012, the team, in collaboration with Labri/Univ. Bordeaux I, has elaborated and experimented a robotic experimental set-up called “Ergo-Robots/FLOWERS Fields” 13. This set-up is not only a way to share our scientific investigations with the general public, but attacks a very important technological challenge impacting the science of developmental robotics: How to design a robot learning experiment that can run continuously and autonomously for several months? Indeed, developmental robotics takes life-long learning and development as one of its central objective and object of study, and thus shall require experimental setups that allow robots to run, learn and develop for extended periods of time. Yet, in practice, this has not been possible so far due to the unavailability of platforms adapted at the same time to learning, exploration, easy and versatile reconfiguration, and extended time of experimentation. Most experiments so far in the field have a duration ranging from a few minutes to a few hours. This is an important obstacle for the progress of developmental robotics, which would need experimental set-ups capable of running for several months. This is exactly the challenge explored by the Ergo-Robots installation, which we have approached by using new generations of affordable yet sophisticated and powerful off-the-shelf servomotors (RX Series from Robotis) combined with an adequately designed software and hardware architecture, as well as processes for streamlined maintenance. The experiment is now running for five months, six days a week, in a public exhibition which has strong constraints over periods of functioning and no continual presence of dedicated technicians/engineers on site. The experiment involved five robots, each with 6 degrees of freedoms, which are endowed with curiosity-driven learning mechanisms allowing them to explore and learn how to manipulate physical objects around them as well as to discover and explore vocal interactions with humans/the visitors. The robots are also playing language games allowing them to invent their own linguistic conventions. A battery of measures has been set up in order to study the evolution of the platform, with the aim of using the results (to be described in an article) as a reference for building future robot learning experiments on extended periods of time, both within the team and in the developmental robotics community. The system has been running during 5 months, 8 hours a day, with no major problems. During the two first months, the platform worked during 390h21mn, and was only stopped during 24h59mn (6 percent of time). After retuning the system based on what we learnt in the two first months, this performance was increased in the three last months: the platform worked for 618h23mn and was only stopped during 17h56mn (2.9 percent of time).
More information available at: http://flowers.inria.fr/ergo-robots.php and http://fondation.cartier.com/.

5.6.2.1. The Ergo-Robots Hardware Platform

Participants: Jerome Bechu [correspondant], Fabien Benureau, Haylee Fogg, Hugo Gimbert, Matthieu Lapeyre, Olivier Ly, Olivier Mangin, Pierre-Yves Oudeyer, Pierre Rouanet.

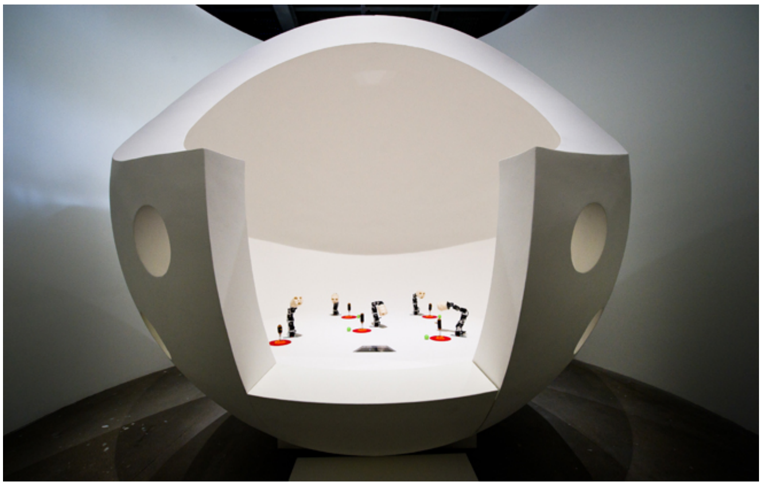


Figure 13. The Ergo-Robot experiment: robot learning experiment running continuously for 5 months at Fondation Cartier pour l'Art Contemporain, exhibition "Mathématiques: Un Dépaysement Soudain".

ErgoRobots 13 is a hardware platform for showcasing a number of curiosity and learning behaviours for the public to interact with. It was designed by the Flowers team in collaboration with Labri/Univ. Bordeaux I. The platform can also have future uses inside the lab for experiments that require more than one robot to complete. Although this system is entirely new this year, a very different previous version existed with the name FLOWERSField. It consists of five ErgoRobots, a control system, an interaction system, a display system, a sound system and a light system. There is an external system which monitors the ErgoRobots which contains a control system, a power system, a surveillance system and a metric capture system. The system has been running during 5 months, 8 hours a day, with no major problems. During the two first months, the platform worked during 390h21mn, and was only stopped during 24h59mn (6 percent of time). After retuning the system based on what we learnt in the two first months, this performance was increased in the three last months: the platform worked for 618h23mn and was only stopped during 17h56mn (2.9 percent of time).

The Ergo-Robot system: The robots themselves are each composed of six motors (see figure). Currently, the heads of the robots have been created in wax by David Lynch and the entire system is displayed at Fondation Cartier inside a large egg shaped orb as shown in the following diagram. The control system module contains both an MMNET1002 control board with an UART-RS485 breakout board which communicates with a ubuntu Linux PC via an ethernet cable. The mment board communicates with the motors, but all other ErgoRobot systems communicate with the PC directly. The sound system is currently externally provided and communicates with the PC. The light system is a series of two or three BlinkM RGB leds placed inside each ErgoRobot head that are controlled through two LinkM USB devices directly with the computer. A kinect placed in front of the system operates as the means for the public to interact with the platform and communicates directly through USB to the PC. The display system is currently an externally provided projector that projects visualisations of the field's current state behind the ErgoRobots.

The external system: This system allows anyone that is monitoring the system to externally control the ErgoRobots system. The PC with which the software control takes place is a Ubuntu Linux system which communicates with the ErgoRobot control system via an ethernet cable. The ErgoRobot harware system can be managed by an external power system which includes a 15.5V bench top power supply for the ErgoRobot motors, an external 12V plug in adapter for the mment board, an external 5V plug in adapter for the LED lights which are all controlled via an emergency stop button. The maintenance system can be located out of direct view of the ErgoRobot field as it has a surveillance system: a kinect that can display the current state of the field. More surveillance is conducted through a metric capture system that communicates with the ErgoRobots to obtain various state values of the ErgoRobots through the motor sensors and other data. This surveillance is not entirely in place as of 2011 and will be implemented in early 2012.

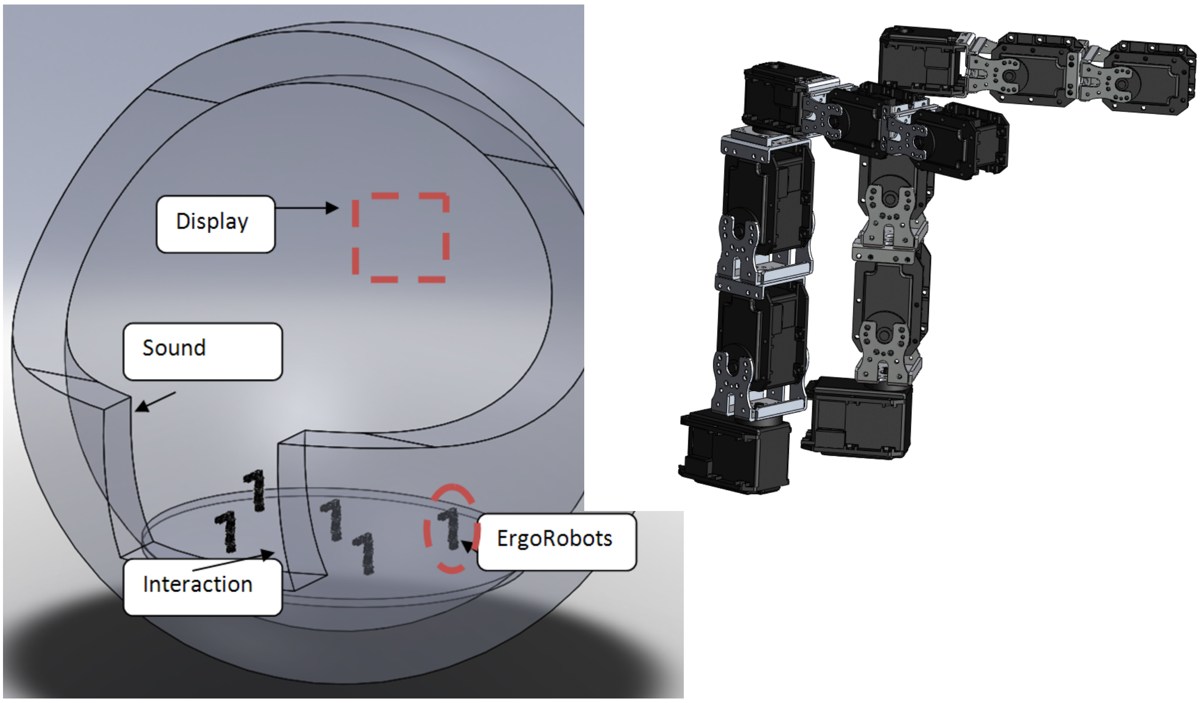


Figure 14. Ergo-Robots

5.6.2.2. Stem Platform for Affordances

Participant: Fabien Benureau [correspondant].

The Stem Platform for Affordances (figure 15 is a hardware platform that is intended for use in the lab for experiments. It features a 6 DOFs arm robot identical to the other robot stems present in the lab, and a physical platform intended for the interaction with objects. Our affordance experiments involves a lot of trials; there was the need for a platform that could reset itself after the robot interacted, as it is an assumption underlying our current algorithms. The stem platform provides exactly that, with the object position and orientation being reseted by the platform autonomously and in less than 10 seconds. This provides the potential to do more than 2000 independent interactions with an object over the course of 12 hours.

The platform also provides sensory capabilities, being able to track the position and orientation of the object at all time. On the hardware side, a camera is used. We investigated both a standard PSEye, that provides a high framerate (120Hz) with noise, and a high quality, firewire camera with professional optics, providing higher resolution, low noise at the expense of a low framerate (15Hz). The latter was kindly provided by the Potioc team. On the software side, computing the position is done by the open-source augmented reality library ARToolkitPlus. On the objects themselves, AR tags are placed.

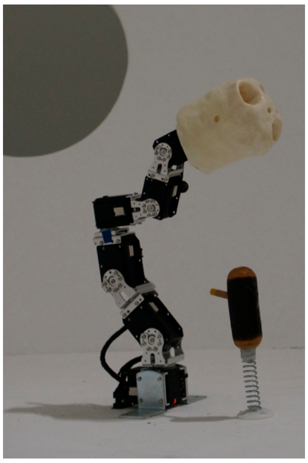


Figure 15. The Stem Platform

The platform is supported by a simulation that reproduce the setup in V-Rep. In order to be able to use the same algorithms for both the hardware and the simulation, a low-level interface was written for Pypot and V-Rep, using the work done by Paul Fudal on V-Rep Bridge.

The complete platform took roughly 3 weeks to make, with 3 additional weeks for the software. The team recently acquired material that would make possible to build a similar platform faster and in a more robust material (wood is used in the first platform). This platform, backed up by its simulation, will allow us to perform planned experiments in a reliable and statistically significant manner.

5.6.2.3. Humanoid Robot Torso

Participant: Haylee Fogg [correspondant].

The Humanoid Robot Torso is a hardware platform that is intended for use in the lab for either experiments or demonstrations 16. It consists of a humanoid robot that contains just a torso, arms with shoulders and grippers, and head. It is entirely new this year, as a new design has been made, and a skeleton built with 3D printing technologies. The arms with the claws contain seven degrees of freedom (including 'grip'). The head consists of a smartphone for the face and an associated camera for the 'eyes' with the ability to move in two degrees (pitch and roll). The hardware is both robotis Dynamixel RX-28 and R-64 motors attached together with standard robotis frames and 3D printed limbs. A wiki has been built, documenting both the hardware and software platform.

5.6.2.4. NoFish platform

Participants: Mai Nguyen [correspondant], Paul Fudal [correspondant], Jérôme Béchu.

The NoFish platform is a hardware platform that is intended for use in the lab for experiments. It consists of an ErgoRobot with an attached fishing rod. The robot is fixed on a table and has in front of him a delimited area where to throw the fishing cap. This area is covered by a camera in order to track the fishing cap and to give its coordinates. The robot is managed by a software written using the Urbi framework. This program controls the robot using pre-programmed moves and also gives a way to uses the robot joint by joint. A second software written in C++ using OpenCV framework tracks the position of the fishing cap and sends the coordinates to the Urbi software controlling the robot. Finally, at the upper layer of the software architecture, MatLab is used to implement different learning algorithms. All MatLab code is able to receive informations from the Urbi part of the software (fishing cap coordinates, joints informations, etc) and also to send order to the robot (position joint by joint, preprogrammed moves, etc). To finish, and because the platform can run a learning algorithms

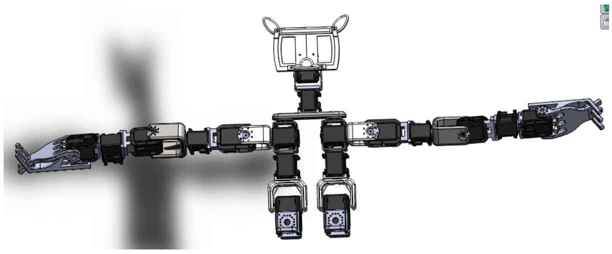


Figure 16. The Humanoid Robot Torso Platform

during a long time, an electric plug managed by the Urbi part of the software is added to the platform to shutdown the power if the robot is blocked or does not respond anymore.

6. New Results

6.1. Autonomous and Social Skill Learning and Development

6.1.1. Active Learning and Intrinsic Motivation

6.1.1.1. Active Learning of Inverse Models with Goal Babbling

Participants: Adrien Baranes, Pierre-Yves Oudeyer.

We have continued to elaborate and study our Self-Adaptive Goal Generation - Robust Intelligent Adaptive Curiosity (SAGG-RIAC) architecture as an intrinsically motivated goal exploration mechanism which allows active learning of inverse models in high-dimensional redundant robots. Based on active goal babbling, this allows a robot to efficiently and actively learn distributions of parameterized motor skills/policies that solve a corresponding distribution of parameterized tasks/goals. The architecture makes the robot sample actively novel parameterized tasks in the task space, based on a measure of competence progress, each of which triggers low-level goal-directed learning of the motor policy parameters that allow to solve it. For both learning and generalization, the system leverages regression techniques which allow to infer the motor policy parameters corresponding to a given novel parameterized task, and based on the previously learnt correspondences between policy and task parameters.
We have conducted experiments with high-dimensional continuous sensorimotor spaces in three different robotic setups: 1) learning the inverse kinematics in a highly-redundant robotic arm, 2) learning omnidirectional locomotion with motor primitives in a quadruped robot 1718, 3) an arm learning to control a fishing rod with a flexible wire. We show that 1) exploration in the task space can be a lot faster than exploration in the actuator space for learning inverse models in redundant robots; 2) selecting goals maximizing competence progress creates developmental trajectories driving the robot to progressively focus on tasks of increasing complexity and is statistically significantly more efficient than selecting tasks randomly, as well as more efficient than different standard active motor babbling methods; 3) this architecture allows the robot to actively discover which parts of its task space it can learn to reach and which part it cannot.
This work was published in the journal Robotics and Autonomous Systems [22].

6.1.1.2. Exploration in Model-based Reinforcement Learning

Participants: Manuel Lopes, Tobias Lang, Marc Toussaint, Todd Hester, Peter Stone, Pierre-Yves Oudeyer.

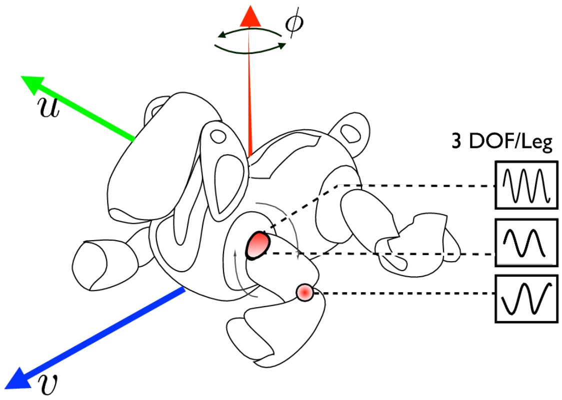


Figure 17. Experimenting SAGG-RIAC for learning an inverse model for omnidirectional locomotion of a quadruped robot. The quadruped robot is controlled using 24 dimensional motor synergies parameterized with 24 continuous values : 12 for the amplitudes and 12 others for the phases of a sinusoid tracked by each motor. Experiments consider a task space u, v, α which corresponds to the 2D position and orientation of the quadruped.

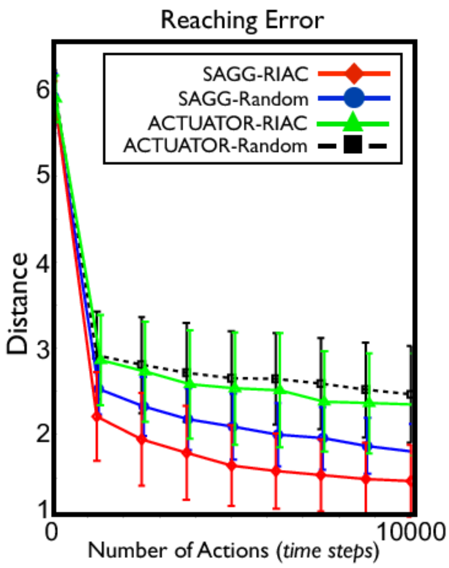
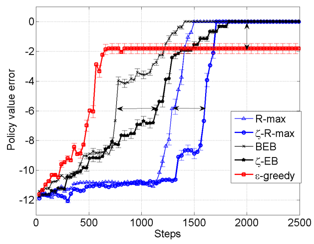
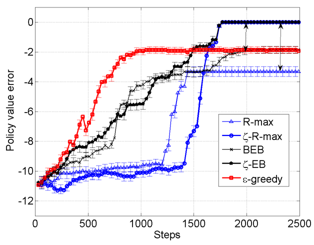


Figure 18. Evolution of the quality of the learnt inverse model for the quadruped robot experiment, depending on various exploration strategies (measured as mean error over a set of uniformly distributed goals generated independantly from learning trials).

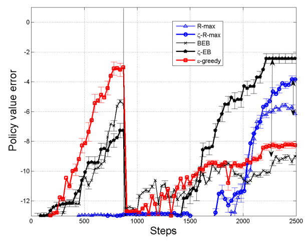
Formal exploration approaches in model-based reinforcement learning estimate the accuracy of the currently learned model without consideration of the empirical prediction error. For example, PAC-MDP approaches such as R-MAX base their model certainty on the amount of collected data, while Bayesian approaches assume a prior over the transition dynamics. We propose extensions to such approaches which drive exploration solely based on empirical estimates of the learner's accuracy and learning progress. We provide a ”sanity check” theoretical analysis, discussing the behavior of our extensions in the standard stationary finite state-action case. We then provide experimental studies demonstrating the robustness of these exploration measures in cases of non-stationary environments or where original approaches are misled by wrong domain assumptions. [46]. Furthermore, we studied how different exploration algorithms can be combine and selected at runtime. Typically the user must hand-tune exploration parameters for each different domain and/or algorithm that they are using. We introduced an algorithm called leo for learning to select among different exploration strategies on-line. This algorithm makes use of bandit-type algorithms to adaptively select exploration strategies based on the rewards received when following them. We show empirically that this method performs well across a set of five domains In contrast, for a given algorithm, no set of parameters is best across all domains. Our results demonstrate that the leo algorithm successfully learns the best exploration strategies on-line, increasing the received reward over static parameterizations of exploration and reducing the need for hand-tuning exploration parameters [42].



(a) *Experiment 1—Correct Assumptions*



(b) *Experiment 2—Violated Assumptions*



(c) *Experiment 3—Change in Dynamics*

Figure 19. Experiments: (a) Like Rmax and BEB with correct assumptions, our algorithms ζ-Rmax and ζ-EB based on an empirical estimation of the learning progress converge to the optimal policy without relying on these assumptions, but take a small extra amount of time. (b) When their assumptions are violated, Rmax and BEB fail to converge, while ζ-Rmax and ζ-EB don't rely on these assumptions and again find the optimal policy. (c) In contrast to existing methods, ζ-Rmax and ζ-EB can cope with the change in transition dynamics after 900 steps and refocus their exploration.

6.1.1.3. The Strategic Student Approach for Life-Long Exploration and Learning

Participants: Manuel LOPES, Pierre-Yves OUDEYER.

We introduced and formalized a general class of learning problems for which a developmental learning strategy is shown to be optimal. This class of problems can be explained using the strategic student metaphor: a student has to learn a number of topics (or tasks) to maximize its mean score, and has to choose strategically how to allocate its time among the topics and/or which learning method to use for a given topic. We show that if the performance curves are sub-modular, then a strategy where time allocation or learning method are chosen in a developmental manner is optimal. We argue that this optimal developmental trajectory can be automatically generated by greedy maximization of learning progress. This optimal strategy amounts to creating a structured developmental exploration where typically easy tasks are first explored, and then progressively more complicated ones are explored. Furthermore, this result holds independently of the nature of the topics and the learning methods used. Then, we show an algorithm, based on multi-armed bandit techniques, that allows empirical online evaluation of learning progress and approximates the optimal solution. Finally, we show that the strategic student problem formulation allows to view in a common framework many previous approaches to active and developmental learning [47].

6.1.1.4. Active Inverse Reinforcement Learning through Generalized Binary Search

Participants: Manuel Lopes, Francisco Melo.

We contributed the first aggressive active learning algorithm for nonseparable multi-class classification. We generalize an existing active learning algorithm for binary classification [107] to the multi-class setting, and identify mild conditions under which the proposed method provably retains the main properties of the original algorithm, namely consistency and sample complexity. In particular, we show that, in the binary case, our method reduces to the original algorithm of [107]. We then contribute an extension of our method to multi-label settings, identify its main properties and discuss richer querying strategies. We conclude the paper with two illustrative application examples. The first application features a standard text-classification problem. The second application scenario features a learning from demonstration setting. In both cases we demonstrate the advantage of our active sampling approach against random sampling. We also discuss the performance of the proposed approach in terms of the derived theoretical bounds.

6.1.1.5. Towards high-dimensional and cumulative task space active exploration

Participant: Benureau Fabien.

One direction of research of the team has been on intrinsic motivation in the context of autonomous learning. Building on the PhD work of Adrien Baranes, the efforts have concentrated on creating algorithms capable to handle high-dimensional spaces and manage context with multiple tasks. The goal is for the learner to be able to autonomously create collection of reusable skills. In this context, two main research efforts have been led this year.

A typical robot is made of chains of joints. We can take advantage of the fact that joints earlier in the chain have more impact that joints further down. Given sensory feedback on the middle of the chain, an algorithm can use this information to boost learning speed and divide the learning space in subsets of smaller dimensions. We wanted to adapt this idea to high dimensional space, and specifically to the interaction with objects; a robotic arm that has already learned an inverse model of its kinematic could reuse this knowledge learn about the mapping between the position of the end-effector and the displacement of an object it is manipulating. Experiments were conducted, but they lead to the conclusion that such an approach, while effective in some specific setting, relies too heavily on a good representation of the end effector position and motion, which, in some cases, requires sensory space of higher dimension that the motor space, thus defeating the purpose. This approach was not found to be robust enough for the type of robotic context our lab is pursuing.

The SAGG-RIAC architecture is an efficient but complex architecture which implementation cannot be easily summarized in a few lines of pseudo-code. This is problematic because it reduces the ability of other research groups to implement and reuse our algorithms for their own work. An effort was started this year to create an implementation of SAGG-RIAC that would be more robust and simpler. The main idea was to use kernels rather than bins to estimate in interest in SAGG-RIAC. This approach led to very promising results, notably in its ability to handle unbounded sensory spaces. We aim at publishing the result of this work in 2013, together with a publicly available implementation of our algorithms with easy to run examples for dissemination of active learning architectures elaborated in the team. This work will also be reused in the participation of the lab into the MaCSi project.

6.1.2. Learning and optimization of motor policies

6.1.2.1. Off-Policy Actor-Critic

Participants: Thomas Degris, Martha White, Richard Sutton.

Actor–critic architectures are an interesting candidate for learning with robots: they can represent complex stochastic policies suitable for robots, they can learn online and incrementally and their per-time-step complexity scales linearly with the number of learned weights. Moreover, interesting connections have been identified in the existing literature with neuroscience. Until recently, however, practical actor–critic methods have been restricted to the on-policy setting, in which the agent learns only about the policy it is executing.

In an off-policy setting, on the other hand, an agent learns about a policy or policies different from the one it is executing. Off-policy methods have a wider range of applications and learning possibilities. Unlike on-policy methods, off-policy methods are able to, for example, learn about an optimal policy while executing an exploratory policy, learn from demonstration, and learn multiple tasks in parallel from a single sensory-motor interaction with an environment. Because of this generality, off-policy methods are of great interest in many application domains.

We have presented the first actor-critic algorithm for off-policy reinforcement learning. Our algorithm is online and incremental, and its per-time-step complexity scales linearly with the number of learned weights. We have derived an incremental, linear time and space complexity algorithm that includes eligibility traces and empirically show better or comparable performance to existing algorithms on standard reinforcement-learning benchmark problems. This work was presented by Degris et al. [38] and was reproduced independently by Saminda Abeyruwan from the University of Miami.

6.1.2.2. Auto-Actor Critic

Participant: Thomas Degris.

As mentioned above, actor–critic architectures are an interesting candidate for robots to learn new skills in unknown and changing environments. However, existing actor–critic architectures, as many machine learning algorithms, require manual tuning of different parameters to work in the real world. To be able to systematize and scale-up skill learning on a robot, learning algorithms need to be robust to their parameters. The Flowers team has been working on making existing actor–critic algorithms more robust to make them suitable to a robotic setting. Results on standard reinforcement learning benchmarks are encouraging. This work will be submitted to international conference related with reinforcement learning. Interestingly, the methods developed in this work also offer a new formalism to think about different existing themes of Flowers research such as curiosity and maturational constraints.

6.1.2.3. Relationship between Black-Box Optimization and Reinforcement Learning

Participant: Freek Stulp.

Policy improvement methods seek to optimize the parameters of a policy with respect to a utility function. There are two main approaches to performing this optimization: reinforcement learning (RL) and black-box optimization (BBO). In recent years, benchmark comparisons between RL and BBO have been made, and there has been several attempts to specify which approach works best for which types of problem classes.

We have made several contributions to this line of research by: 1) Defining four algorithmic properties that further clarify the relationship between RL and BBO. 2) Showing how the derivation of ever more powerful RL algorithms displays a trend towards BBO. 3) Continuing this trend by applying two modifications to the state-of-the-art PI^2 algorithm, which yields an algorithm we denote PI^{BB}. We show that PI^{BB} is a BBO algorithm, and, more specifically, that it is a special case of the state-of-the-art CMAES algorithm. 4) Demonstrating that the simpler PI^{BB} achieves similar or better performance than PI^2 on several evaluation tasks. 5) Analyzing why BBO outperforms RL on these tasks. These contributions have been published on HAL [69], and have been submitted to JMLR.

This work has also resulted in the novel PI^2-CMA, PI^2-CMAES algorithms, which are presented in [63], [60], [62]

6.1.2.4. Reinforcement Learning with Sequences of Motion Primitives for Robust Manipulation

Participant: Freek Stulp.

Physical contact events often allow a natural decomposition of manipulation tasks into action phases and subgoals. Within the motion primitive paradigm, each action phase corresponds to a motion primitive, and the subgoals correspond to the goal parameters of these primitives. Current state-of-the-art reinforcement learning algorithms are able to efficiently and robustly optimize the parameters of motion primitives in very high-dimensional problems. These algorithms often consider only shape parameters, which determine the trajectory between the start- and end-point of the movement. In manipulation, however, it is also crucial to optimize the goal parameters, which represent the subgoals between the motion primitives. We therefore extend the policy improvement with path integrals (PI^2) algorithm to simultaneously optimize shape and goal parameters. Applying simultaneous shape and goal learning to sequences of motion primitives leads to the novel algorithm PI^2-Seq. We use our methods to address a fundamental challenge in manipulation: improving the robustness of everyday pick-and-place tasks. This work was published in IEEE Transactions on Robotics [31] and Robotics and Autonomous Systems [26].

6.1.2.5. Model-free Reinforcement Learning of Impedance Control in Stochastic Environments

Participant: Freek Stulp.

For humans and robots, variable impedance control is an essential component for ensuring robust and safe physical interaction with the environment. Humans learn to adapt their impedance to specific tasks and environments; a capability which we continually develop and improve until we are well into our twenties. We have reproduced functionally interesting aspects of learning impedance control in humans on a simulated robot platform.

As demonstrated in numerous force field tasks, humans combine two strategies to adapt their impedance to perturbations, thereby minimizing position error and energy consumption: 1) if perturbations are unpredictable, subjects increase their impedance through co-contraction; 2) if perturbations are predictable, subjects learn a feed-forward command to offset the perturbation. We show how a 7-DOF simulated robot demonstrates similar behavior with our model-free reinforcement learning algorithm , by applying deterministic and stochastic force fields to the robot's end-effector. We show the qualitative similarity between the robot and human movements.

Our results provide a biologically plausible approach to learning appropriate impedances purely from experience, without requiring a model of either body or environment dynamics. Not requiring models also facilitates autonomous development for robots, as pre-specified models cannot be provided for each environment a robot might encounter. This work was published in IEEE Transactions on Autonomous Mental Development [29].

6.1.2.6. Probabilistic optimal control: a quasimetric approach

Participants: Clément Moulin-Frier, Jacques Droulez, Steve Nguyen.

During his previous post-doc at the Laboratoire de Physiologie de la Perception et de l'Action (Collège de France, Paris), Clément Moulin-Frier joined Jacques Droulez and Steve N'Guyen to work on an alternative and original approach of probabilistic optimal control called the quasimetric. A journal paper (soon to be submitted) was written in 2012, where the authors propose a new approach for dealing with control under uncertainty.

6.1.3. Social learning and intrinsic motivation

6.1.3.1. Optimal Teaching on Sequential Decision Tasks

Participants: Manuel Lopes, Maya Cakmak.

A helpful teacher can significantly improve the learning rate of an autonomous learning agent. Teaching algorithms have been formally studied within the field of Algorithmic Teaching. These give important insights into how a teacher can select the most informative examples while teaching a new concept. However the field has so far focused purely on classification tasks. We introduced a novel method for optimally teaching sequential decision tasks. We present an algorithm that automatically selects the set of most informative demonstrations and evaluate it on several navigation tasks. Next, we present a set of human subject studies that investigate the optimality of human teaching in these tasks. We evaluate examples naturally chosen by human teachers and found that humans are generally sub-optimal. Then based on our proposed optimal teaching algorithm we try to elicit better teaching from humans. We do this by explaining the intuition of the teaching algorithm in an informal language prior to the teaching task. We found that this improves the examples elicited from human teachers on all considered tasks. This shows that a simple modification the instructions given to human teachers, has the potential of greatly improving the performance of the agent trained by the human [32].

6.1.3.2. Socially Guided Intrinsic Motivation for Skill Learning

Participants: Sao Mai Nguyen, Pierre-Yves Oudeyer.

We have explored how social interaction can bootstrap the learning of a robot for motor learning. We first studied how simple demonstrations by teachers could have a bootstrapping effect on autonomous exploration with intrinsic motivation by building a learner who uses both imitation learning and SAGG-RIAC algorithm [22], and thus designed the SGIM-D (Socially Guided Intrinsic Motivation by Demonstration) algorithm [105]. We then investigated on the reasons of this bootstrapping effect [55], to show that demonstrations by teachers can both enhance more tasks to be explored, as well as favor more easily generalized actions to be used. This analysis is generalizable for all algorithms using social guidance and goal-oriented exploration. We then proposed to build a strategic learner who can learn multiple tasks and with multiple strategies. An overview and theoretical study of multi-task, multi-strategy Strategic Learning is presented in [47]. We also forsook to build a learning algorithm for more natural interaction with the human users. We first designed the SGIM-IM algorithm so that it can determine itself when it should ask for help from the teacher while trying to explore autonomously as long as possible so as to use as little of the teacher's time as possible [54]. After tackling with the problem of how and when to learn, we also investigated an active learner who can determine who to ask for help: in the case of two teachers available, SGIM-IM can determine which strategy to adopt between autonomous exploration and learning by demonstration, and which teacher enhances most learning progress for the learner [56], and ask him for help.

While the above results have been shown in simulation environments: of a simple deterministic air hockey game (fig. 20), and a stochastic fishing experiment with a real-time physical simulator (fig. 21), we are now building the experimental setup of the fishing experiment in order to carry out the experiments with naive users.

6.1.3.3. Adaptive task execution for implicit human-robot coordination

Participants: Ievgen Perederieiev, Manuel Lopes, Freek Stulp.

We began a project which goal is to study how computational models of multi-agent systems can be applied in situations where one agent is a human. We aim at applications where robots collaborate with humans for achieving complex tasks..

A very important capability for efficient collaborative work is the mutual agreement of a task and the ability to predict the behavior of others. We address such aspect by studying methods that increase the predictability of the robot actions. An efficient motor execution becomes the one that not just optimize speed and minimizes energy but also the one that improves the reliability of the team behavior. We are studying policy gradient methods and working on policy improvement algorithms (PI^2, CEM and $CMAES$). A feasibility study will consider a simple task between a robot and a person where the goal is to coordinate the way a set of three colored buttons is pressed.

6.1.3.4. Formalizing Imitation Learning

Participants: Thomas Cederborg, Pierre-Yves Oudeyer.

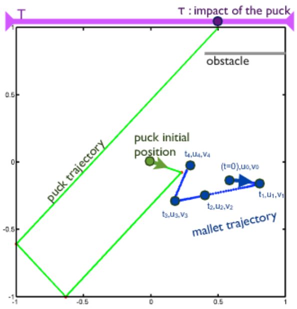


Figure 20. Illustration of SGIM-D and SGIM-IM algorithms

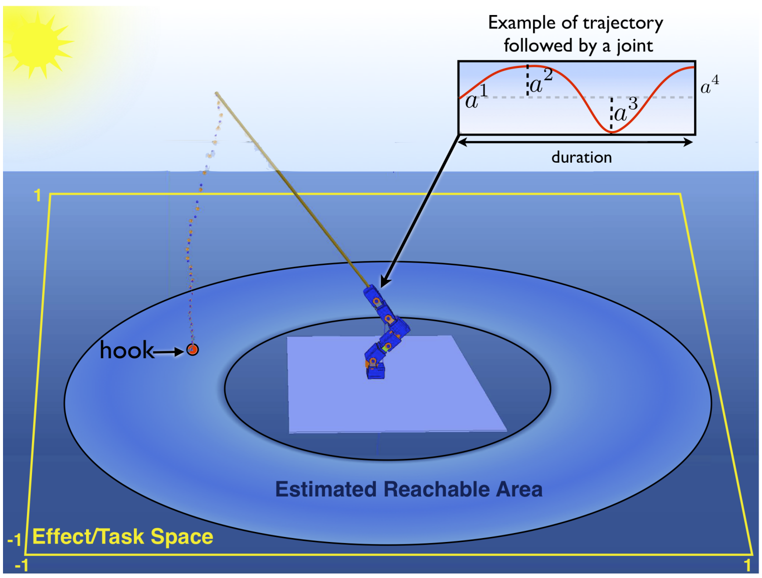


Figure 21. Illustration of SGIM-D and SGIM-IM algorithms

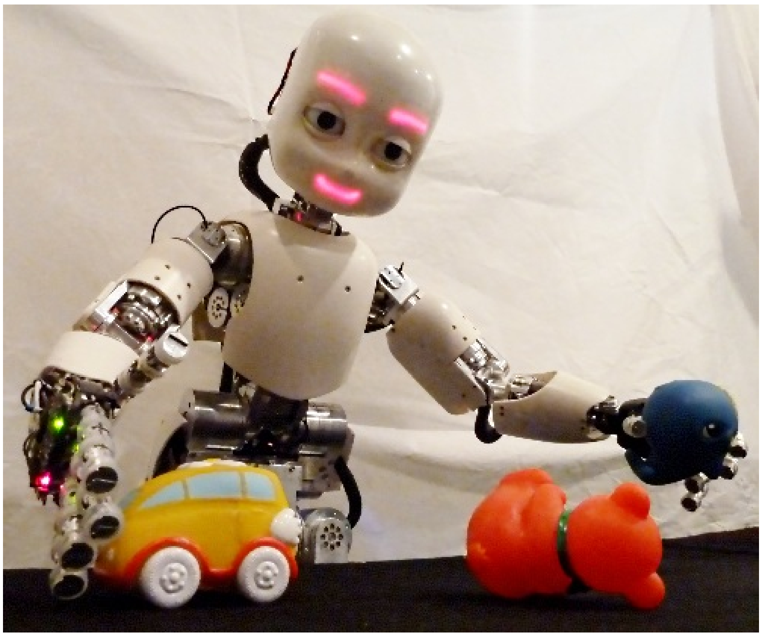


Figure 22. Illustration of SGIM-D and SGIM-IM algorithms

An original formalization of imitation learning was elaborated. Previous attempts to systematize imitation learning has been limited to categorizing different types of demonstrator goals (for example defining success in terms of the sequential joint positions of a dance, or in terms of environmental end states), and/or been limited to a smaller subset of imitation (such as learning from tele-operated demonstrations). The formalism proposed attempts to describe a large number of different types of learning algorithms using the same notation. Any type of algorithm that modifies a policy based on observations of a human, is treated as an interpretation hypothesis of this behavior. One example would be an update algorithm that updates a policy, partially based on the hypothesis that the demonstrator succeeds at demonstrations with probability 0.8, or an update algorithm that assumes that a scalar value is an accurate evaluation of an action compared to the latest seven actions. The formalism aims to give a principled way of updating these hypotheses, either rejecting some of a set of hypotheses regarding the same type of behavior, or set of parameters of an hypothesis. Any learning algorithm that modifies policy based on observations of a human that wants an agent to do something or act in some way, is describable as an interpretation hypothesis. If the learning algorithm is static, this simply corresponds to an hypothesis that is not updated based on observations. A journal article is currently being written.

6.1.4. Unsupervised learning of motor primitives

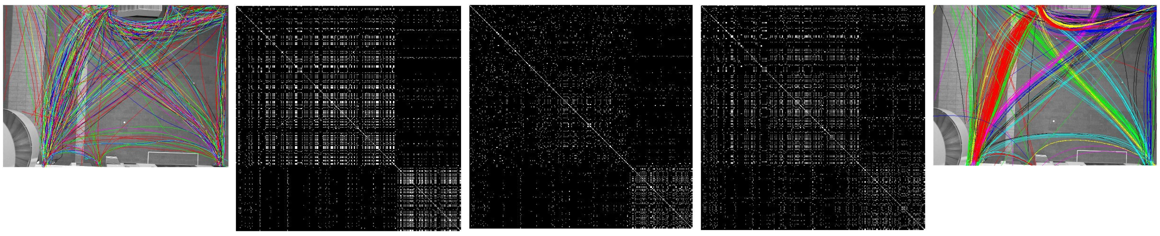
6.1.4.1. Clustering activities

Participants: Manuel Lopes, Luis Montesano.

Learning behaviors from data has applications in surveillance and monitoring systems, virtual agents and robotics among others. In our approach, ww assume that in a given unlabeled dataset of multiple behaviors, it is possible to find a latent representation in a controller space that allows to generate the different behaviors. Therefore, a natural way to group these behaviors is to search a common control system that generate them accurately.

Clustering behaviors in a latent controller space has two major challenges. First, it is necessary to select the control space that generate behaviors. This space will be parameterized by a set of features that will change for different behaviors. Usually, each controller will minimize a cost function with respect to several task features. The latent representation is in turn defined by the selected features and their corresponding weight. Second, an unknown number of such controllers is required to generate different behaviors and the grouping must be based on the ability of the controller to generate the demonstrations using a compact set of controllers.

We propose a Dirichlet Process based algorithm to cluster behaviors in a latent controller space which encodes the dynamical system generating the observed trajectories. The controller uses a potential function generated as a linear combination of features. To enforce sparsity and automatically select features for each cluster independently, we impose a conditional Laplace prior over the controller parameters. Based on this models, we derive a sparse Dirichlet Process Mixture Model (DPMM) algorithm that estimates the number of behaviors and a sparse latent controller for each of them based on a large set of features.



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

Figure 23. EIFPD dataset. (a) Trajectories of the EIFPD to be clustered (color is non-informative). (b-d) correspondence matrix for the 474 trajectories for the labeled ground truth, the KMeans in measurement space and the DPMM, respectively. (e) Reconstructed trajectories from the initial point using the estimated parameters of the DPMM algorithm. Due to the large number of clusters (37), colors are repeated for different clusters.

6.1.4.2. Learning the Combinatorial Structure of Demonstrated Behaviors with Inverse Feedback Control

Participants: Olivier Mangin, Pierre-Yves Oudeyer.

We have elaborated and illustrated a novel approach to learning motor skills from demonstration. This approach combines ideas from inverse feedback learning, in which actions are assumed to solve a task, and dictionary learning. In this work we introduced a new algorithm that is able to learn behaviors by assuming that the observed complex motions can be represented in a smaller dictionary of concurrent tasks. We developed an optimization formalism and show how we can learn simultaneously the dictionary and the mixture coefficients that represent each demonstration. We presented results on a idealized model where a set of potential functions represents human objectives or preferences for achieving a task in [51].

6.1.5. Maturational learning

6.1.5.1. Emergent Proximo-Distal Maturation through Adaptive Exploration

Participants: Freek Stulp, Pierre-Yves Oudeyer.

Life-long robot learning in the high-dimensional real world requires guided and structured exploration mechanisms. In this developmental context, we have investigated the use of the PI^2-CMAES episodic reinforcement learning algorithm, which is able to learn high-dimensional motor tasks through adaptive control of exploration. By studying PI^2-CMAES in a reaching task on a simulated arm, we observe two developmental properties. First, we show how PI^2-CMAES autonomously and continuously tunes the global exploration/exploitation trade-off, allowing it to re-adapt to changing tasks. Second, we show how PI^2-CMAES spontaneously self-organizes a maturational structure whilst exploring the degrees-of-freedom

(DOFs) of the motor space. In particular, it automatically demonstrates the so-called *proximo-distal maturation* observed in humans: after first freezing distal DOFs while exploring predominantly the most proximal DOF, it progressively frees exploration in DOFs along the proximo-distal body axis. These emergent properties suggest the use of PI^2-CMAES as a general tool for studying reinforcement learning of skills in life-long developmental learning contexts. This work was published in the IEEE International Conference on Developement and Learning [60].

6.1.5.2. Interaction of Maturation and Intrinsic Motivation for Developmental Learning of Motor Skills in Robots

Participants: Adrien Baranes, Pierre-Yves Oudeyer.

We have introduced an algorithmic architecture that couples adaptively models of intrinsic motivation and physiological maturation for autonomous robot learning of new motor skills. Intrinsic motivation, also called curiosity-driven learning, is a mechanism for driving exploration in active learning. Maturation denotes here mechanisms that control the evolution of certain properties of the body during development, such as the number and the spatio-temporal resolution of available sensorimotor channels. We argue that it is useful to introduce and conceptualize complex bidirectional interactions among these two mechanisms, allowing to actively control the growth of complexity in motor development in order to guide efficiently exploration and learning. We introduced a model of maturational processes, taking some functional inspiration from the myelination process in humans, and show how it can be coupled in an original and adaptive manner with the intrinsic motivation architecture SAGG-RIAC (Self-Adaptive Goal Generation - Robust Intelligent Adaptive Curiosity algorithm), creating a new system, called McSAGG-RIAC. We then conducted experiments to evaluate both qualitative and quantitative properties of these systems when applied to learning to control a high-dimensional robotic arm, as well as to learning omnidirectional locomotion in a quadruped robot equipped with motor synergies. We showed that the combination of active and maturational learning can allow to gain orders of magnitude in learning speed as well as reach better generalization performances. A journal article is currently being written.

6.1.6. Morphological computation and body intelligence

6.1.6.1. Comparative Study of the Role of Trunk in Human and Robot Balance Control

Participants: Matthieu Lapeyre [correspondant], Christophe Halgand, Jean-René Cazalet, Etienne Guillaud, Pierre-Yves Oudeyer.

Numerous studies in the field of functional motor rehabilitation were devoted to understanding the functioning of members but few are interested in the coordination of the trunk muscles and the relationship between axial and appendicular motricity which is essential in maintaining balance during travel. Acquiring new knowledge on this subject is a prerequisite in the development of new therapeutic strategies to restore motor function to the overall development of robotic orthosis that would assist the movement. Many robotic orthosis using EMG signals were unfortunately using few joints [85] and a system for controlling a multi articulated spine has not yet been developed. We propose here to use a multidisciplinary approach to define the neuro-mechanical principles where an axial system is operating in synergy with human and robot limbs.

To bring us a theoretical framework, we chose to study the reactions of the Acroban humanoid robot. Including 5 joints in the trunk, Acroban can reproduce in part the fluid movements of the human body [98] and especially to test its behavior when its trunk is held fixed or his arms are no longer used for rebalance. To disrupt postural balance in humans and robots, we have developed a low cost mobile platform (see Figure 24). This platform is made up of a broad stable support $(0.8x5m)$ mounted on a skateboard having a power of 800W. The substitution of the initial order of skate by an embedded microcontroller allows us to generate mono-axial perturbations precise intensity and duration to ensure repeatability of the disturbance. We capture movements (Optitrack 250Hz) and record the acceleration of the platform (accelerometer embedded 2kHz), the center of pressure (WiiBalanceBoard 60Hz), and electromyography (EMG).

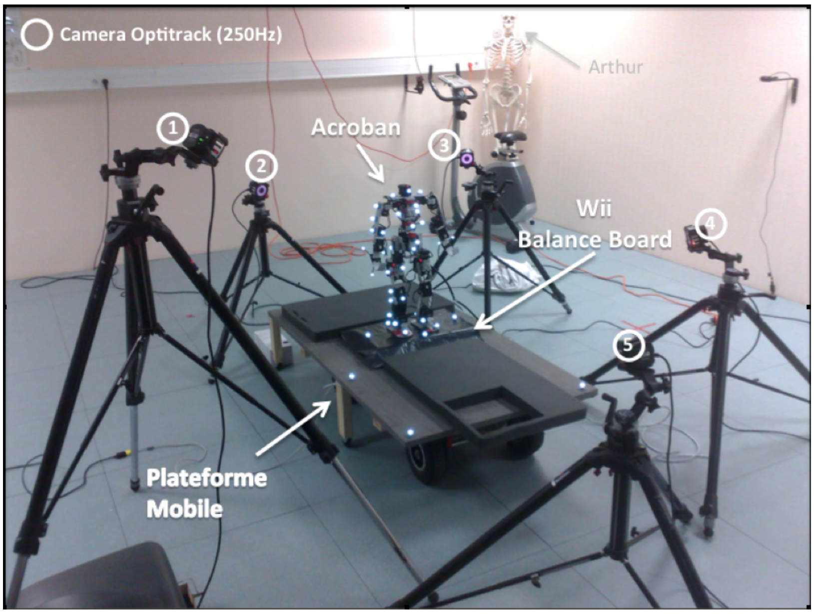


Figure 24. Experimental setup for comparative study of the role of the trunk in human and robot balance control

The experimental device (mobile platform and synchronized recordings) is operational. Preliminary experiments have allowed us to refine the profiles of disturbance on the robot Acroban. The analysis of preliminary results is in progress. Following this study, we hope to improve the modeling of the motor system in humans and robotic simulation as a basis for the development of robotic orthosis axial system. Second, the results provide a basis for improved balancing of Acroban primitives but also the development of future humanoid robots.

6.2. Autonomous and Social Perceptual Learning

6.2.1. The Impact of Human-Robot Interfaces on the Learning of Visual Objects

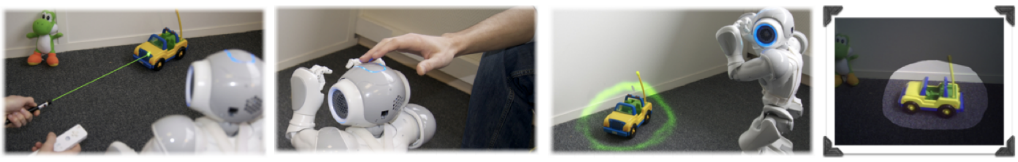
Participants: Pierre Rouanet, Pierre-Yves Oudeyer, Fabien Danieau, David Filliat.

We have continued and finalized a large-scale study of the impact of interfaces allowing non-expert users to efficiently and intuitively teach a robot to recognize new visual objects. We identified challenges that need to be addressed for real-world deployment of robots capable of learning new visual objects in interaction with everyday users. We argue that in addition to robust machine learning and computer vision methods, well-designed interfaces are crucial for learning efficiency. In particular, we argue that interfaces can be key in helping non-expert users to collect good learning examples and thus improve the performance of the overall learning system. Then, we have designed four alternative human-robot interfaces: three are based on the use of a mediating artifact (smartphone, wiimote, wiimote and laser), and one is based on natural human gestures (with a Wizard-of-Oz recognition system). These interfaces mainly vary in the kind of feedback provided to the user, allowing him to understand more or less easily what the robot is perceiving, and thus guide his way of providing training examples differently. We then evaluated the impact of these interfaces, in terms of learning efficiency, usability and user's experience, through a real world and large scale user study. In this experiment, we asked participants to teach a robot twelve different new visual objects in the context of a robotic game. This game happens in a home-like environment and was designed to motivate and engage users in an interaction where using the system was meaningful. We then analyzed results that show significant differences among

interfaces. In particular, we showed that interfaces such as the smartphone interface allows non-expert users to intuitively provide much better training examples to the robot, almost as good as expert users who are trained for this task and aware of the different visual perception and machine learning issues. We also showed that artifact-mediated teaching is significantly more efficient for robot learning, and equally good in terms of usability and user's experience, than teaching thanks to a gesture-based human-like interaction.
This work was accepted for publication in the IEEE Transactions on Robotics [28].

(a) draw the attention toward an object (b) trigger the capture (c) encircle the area of the object (d) the new learning example

Figure 25. Smartphone Interface. To make the robot collect a new learning example, users have to first draw the robot's attention toward the object they want to teach through simple gestures. Once the robot sees the object, they touch the head of the robot to trigger the capture. Then, they directly encircle the area of the image that represents the object on the screen. The selected area is then used as the new learning example. The combination of the video stream and the gestures facilitate the achievement of joint attention.



(a) draw the attention toward an object (b) trigger the capture (c) encircle the area of the object (d) the new learning example

Figure 26. Wiimote + laser pointer interface. With this interface users can draw the robot's attention with a laser pointer toward an object. The laser spot is automatically tracked by the robot. They can ensure that the robot detects the spot thanks to haptic feedback on the Wiimote. Then, they can touch the head of the robot to trigger the capture of a new learning example. Finally, they encircle the object with the laser pointer to delimit its area which will be defined as the new learning example.

6.2.2. Curiosity-driven exploration and interactive learning of visual objects with the ICub robot

Participants: Mai Nguyen, Serena Ivaldi, Natalia Lyubova, Alain Droniou, Damien Gerardeaux-Viret, David Filliat, Vincent Padois, Olivier Sigaud, Pierre-Yves Oudeyer.

We studied how various mechanisms for cognition and learning, such as curiosity, action selection, imitation, visual learning and interaction monitoring, can be integrated in a single embodied cognitive architecture. We have conducted an experiment with the iCub robot for active recognition of objects in 3D through curiosity-driven exploration, in which the robot can manipulate the robot or ask a human user to manipulate objects to gain information and recognise better objects (fig. 22). For this experiment carried out within the MACSi project, we address the problem of learning to recognise objects in a developmental robotics scenario. In a life-long learning perspective, a humanoid robot should be capable of improving its knowledge of objects

Figure 27. The real world environment designed to reproduce a typical living room. Many objects were added in the scene in order to make the environment cluttered.

with active perception. Our approach stems from the cognitive development of infants, exploiting active curiosity-driven manipulation to improve perceptual learning of objects. These functionalities are implemented as perception, control and active exploration modules as part of the Cognitive Architecture of the MACSi project. We integrated a bottom-up vision system based on swift feature points and motor-primitive based robot control with the SGIM-ACTS algorithm (Socially Guided Intrinsic Motivation with Active Choice of Task and Strategy as the active exploration module. SGIM-ACTS is a strategic learner who actively chooses which task to concentrate on, and which strategy is better according to this task. It thus monitors the learning progress for each strategy on all kinds of tasks, and actively interacts with the human teacher. We obtained an active object recognition approach, which exploits curiosity to guide exploration and manipulation, such that the robot can improve its knowledge of objects in an autonomous and efficient way. Experimental results show the effectiveness of our approach: the humanoid iCub is now capable of deciding autonomously which actions must be performed on objects in order to improve its knowledge, requiring a minimal assistance from its caregiver. This work constitutes the base for forthcoming research in autonomous learning of affordances.

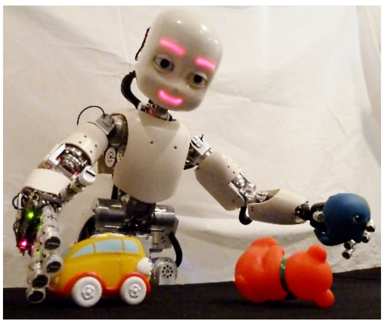


Figure 28. iCub performing curiosity-driven exploration and active recognition of visual objects in 3D

6.2.3. Discovering object concept through developmental learning

Participants: Natalia Lyubova, David Filliat.

The goal of this work is to design a visual system for a humanoid robot. Taking inspiration from child perception and following the principles of developmental robotics, the robot should detect and learn objects from interactions with people and from experiments it performs with objects, avoiding the use of image databases or of a separate training phase. In our model, all knowledge is therefore iteratively acquired from low-level features and builds up hierarchical object models, which are robust to changes in the environment, background and camera motion. In our scenario, people in front of the robot are supposed to interact with objects to encourage the robot to focus on them. We therefore assume that the robot is attracted by motion and we segment possible objects based on clustering of the optical flow. Additionally, the depth information from a Kinect is used to filter visual input, considering the constraints of the robot's working area and to refine the object contours obtained from motion segmentation.

The appearance of objects is encoded following the Bag of Visual Words approach with incremental dictionaries. We combine several complementary features to maximize the completeness of the encoded information (SURF descriptor and superpixels with associated colors) and construct pairs and triples of these features to integrate local geometry information. These features make it possible to decide if the current view has been already seen or not. A multi-view object model is then constructed by associating recognized views and views tracked during manipulations with an object.

This system is implemented on the iCub humanoid robot, which detects objects in the visual space and characterizes their appearance, their relative position and their occurrence statistics. The experiments were performed with up to ten objects; each of them was manipulated by a person during 1-2 minutes. Once the vocabulary reached a sufficient amount of knowledge, the robot was able to reliably recognize most of objects [48], [49], [43].

6.2.4. Unsupervised object categorization

Participants: Natalia Lyubova, David Filliat.

The developed unsupervised algorithm allows to identify segmented units of attention based on motion and depth information (proto-objects) into different categories such as robot hands, objects and humans.

The robot self-body category is discovered from the correlation between the proto-object positions and proprioception on the robot arms. This correlation it estimated by computing the mutual information between the changes in robot motor joints and the motion behavior of proto-objets in the visual field. The arm joints states are recorded from the robot and quantized to a vocabulary of possible arm configurations. The visual space is analyzed at the level of visual clusters that divide the perception field into regular regions. The mutual information is computed from the occurrence probabilities of the arm configurations and visual clusters.

In case of high correlation, the visual cluster is identified as a robot hand. Among the remaining proto-objects, objects are distinguished from human hands based on their quasi-static nature. Since most of objects don't move by themselves but rather are displaced by external forces, the object category is associated with regions of the visual space moving together mostly with recognized robot hands or human parts. This process make it possible to recognize the robot hands, even in case of changing appearance, and to learn to separate objects from parts of the caregivers bodies.

6.2.5. Efficient online bootstrapping of sensory representations

Participant: Alexander Gepperth.

This work [24] is a simulation-based investigation exploring a novel approach to the open-ended formation of multimodal representations in autonomous agents. In particular, we addressed here the issue of transferring (bootstrapping) features selectivities between two modalities, from a previously learned or innate reference representation to a new induced representation. We demonstrated the potential of this algorithm by several experiments with synthetic inputs modeled after a robotics scenario where multimodal object representations are bootstrapped from a (reference) representation of object affordances, focusing particularly on typical challenges in autonomous agents: absence of human supervision, changing environment statistics and limited computing power. We proposed an autonomous and local neural learning algorithm termed PROPRE (projection-prediction) that updates induced representations based on predictability: competitive advantages are given to those feature-sensitive elements that are inferable from activities in the reference representation, the key ingredient being an efficient online measure of predictability controlling learning. We verified that the proposed method is computationally efficient and stable, and that the multimodal transfer of feature selectivity is successful and robust under resource constraints. Furthermore, we successfully demonstrated robustness to noisy reference representations, non-stationary input statistics and uninformative inputs.

6.2.6. Simultaneous concept formation driven by predictability

Participants: Alexander Gepperth, Louis-Charles Caron.

This work [40] was conducted in the context of developmental learning in embodied agents who have multiple data sources (sensors) at their disposal. We developed an online learning method that simultaneously discovers meaningful concepts in the associated processing streams, extending methods such as PCA, SOM or sparse coding to the multimodal case. In addition to the avoidance of redundancies in the concepts derived from single modalities, we claim that meaningful concepts are those who have statistical relations across modalities. This is a reasonable claim because measurements by different sensors often have common cause in the external world and therefore carry correlated information. To capture such cross-modal relations while avoiding redundancy of concepts, we propose a set of interacting self-organization processes which are modulated by local predictability. To validate the fundamental applicability of the method, we conducted a plausible simulation experiment with synthetic data and found that those concepts which are predictable from other modalities successively ”grow”, i.e., become overrepresented, whereas concepts that are not predictable become systematically under-represented. We additionally explored the applicability of the developed method to real-world robotics scenarios.

6.2.7. The contribution of context: a case study of object recognition in an intelligent car

Participants: Alexander Gepperth, Michael Garcia Ortiz.

In this work [23], we explored the potential contribution of multimodal context information to object detection in an ”intelligent car”. The used car platform incorporates subsystems for the detection of objects from local visual patterns, as well as for the estimation of global scene properties (sometimes denoted scene context or just context) such as the shape of the road area or the 3D position of the ground plane. Annotated data recorded on this platform is publicly available as the a ”HRI RoadTraffic” vehicle video dataset, which formed the basis for the investigation. In order to quantify the contribution of context information, we investigated whether it can be used to infer object identity with little or no reference to local patterns of visual appearance. Using a challenging vehicle detection task based on the ”HRI RoadTraffic” dataset, we trained selected algorithms (context models) to estimate object identity from context information alone. In the course of our performance evaluations, we also analyzed the effect of typical real-world conditions (noise, high input dimensionality, environmental variation) on context model performance. As a principal result, we showed that the learning of context models is feasible with all tested algorithms, and that object identity can be estimated from context information with similar accuracy as by relying on local pattern recognition methods. We also found that the use of basis function representations [1] (also known as ”population codes” allows the simplest (and therefore most efficient) learning methods to perform best in the benchmark, suggesting that the use of context is feasible even in systems operating under strong performance constraints.

6.2.8. Co-training of context models for real-time object detection

Participant: Alexander Gepperth.

In this work[41], we developed a simple way to reduce the amount of required training data in context-based models of real- time object detection and demonstrated the feasibility of our approach in a very challenging vehicle detection scenario comprising multiple weather, environment and light conditions such as rain, snow and darkness (night). The investigation is based on a real-time detection system effectively composed of two trainable components: an exhaustive multiscale object detector (signal-driven detection), as well as a module for generating object-specific visual attention (context models) controlling the signal-driven detection process. Both parts of the system require a significant amount of ground-truth data which need to be generated by human annotation in a time-consuming and costly process. Assuming sufficient training examples for signal-based detection, we showed that a co-training step can eliminate the need for separate ground-truth data to train context models. This is achieved by directly training context models with the results of signal-driven detection. We demonstrated that this process is feasible for different qualities of signal-driven detection, and maintains the performance gains from context models. As it is by now widely accepted that signal-driven object detection can be significantly improved by context models, our method allows to train strongly improved detection systems without additional labor, and above all, cost.

6.3. Joint Learning and Development of Language and Action

6.3.1. Learning to recognize parallel motion primitives with linguistic descriptions using Non-Negative Matrix Factorization

Participants: Olivier Mangin, Pierre-Yves Oudeyer.

We have elaborated and experimented a novel approach to joint language and motor learning from demonstration. It enables discovery of a dictionary of gesture and linguistic primitives, that can be combined in parallel to represent training data as well as novel activities in the form of combinations of known gestures. These methods and the results of our experiments participate in addressing two main issues of developmental robotics: 1) symbol grounding for language learning; 2) achieving compositionality in motor-learning from demonstration, which enables re-using knowledge and thus scaling to complex tasks. In particular, we are interested in learning motor primitives active in parallel, a less explored way of combining such primitives. To address these challenges we have explored and studied the use of nonnegative matrix factorization to discover motor primitives from histogram representations of data acquired from real demonstrations of dancing movements. Initial results were presented in [99] and further results are presented in [52].

6.3.2. Curiosity-driven phonetic learning

Participants: Clément Moulin-Frier, Pierre-Yves Oudeyer.

We study how developmental phonetic learning can be guided by pure curiosity-driven exploration, also called intrinsically motivated exploration. Phonetic learning refers here to learning how to control a vocal tract to reach acoustic goals. We compare three different exploration strategies for learning the auditory-motor inverse model: random motor exploration, random goal selection with reaching, and curiosity-driven active goal selection with reaching. Using a realistic vocal tract model, we show how intrinsically motivated learning driven by competence progress can generate automatically developmental structure in both articulatory and auditory modalities, displaying patterns in line with some experimental data from infants. This work has been published in [53] and received the best paper award in computational models of development at the International Conference on Development and Learning, Epirob, San Diego, 2012.

We are now working on applying this approach to the control of a more complex articulatory synthesizer. We are interested in using the free software Praat, a powerful tool allowing to synthesize a speech signal from a trajectory in a 29-dimensional space of respiratory and oro-facial muscles. Numerous acoustic features can in turn be extracted from the synthesized sound, among which the Mel-frequency cepstral coefficients. Our hope is that a developmental robotics approach applied to a realistic articulatory model can appropriately manage the learning process of this complex mapping in high-dimensional spaces , and that observed developmental sequences can lead to interesting experimental data comparisons and predictions. In particular, using such a dynamic model controlled by muscle activity could hopefully allow to relate our results to more common speech acquisition data, in particular regarding infraphonological exploration and babbling.

6.3.3. Towards robots with teleological action and language understanding

Participants: Britta Wrede, Katharina Rohlfing, Jochen Steil, Sebastian Wrede, Jun Tani, Pierre-Yves Oudeyer.

It is generally agreed upon that in order to achieve generalizable learning capabilities of robots they need to be able to acquire compositional structures - whether in language or in action. However, in human development the capability to perceive compositional structure only evolves at a later stage. Before the capability to understand action and language in a structured, compositional way arises, infants learn in a holistic way which enables them to interact in a socially adequate way with their social and physical environment even with very limited understanding of the world, e.g. trying to take part in games without knowing the exact rules. This capability endows them with an action production advantage which elicits corrective feedback from a tutor, thus reducing the search space of possible action interpretations tremendously. In accordance with findings from developmental psychology we argue that this holistic way is in fact a teleological representation encoding a goal-directed perception of actions facilitated through communicational frames. This observation leads to a range of consequences which need to be verified and analysed in further research. We have written an article [64] where we discussed two hypotheses how this can be made accessible for action learning in robots: (1) We explored the idea that the teleological approach allows some kind of highly reduced one shot learning enabling the learner to perform a meaningful, although only partially correct action which can then be further refined through compositional approaches. (2) We discussed the possibility to transfer the concept of "conversational frames" as recurring interaction patterns to the action domain, thus facilitating to understand the meaning of a new action. We conclude that these capabilities need to be combined with more analytical compositional learning methods in order to achieve human-like learning performance.

6.3.4. Imitation Learning and Language

Participants: Thomas Cederborg, Pierre-Yves Oudeyer.

We have studied how context-dependant imitation learning of new skills and language learning could be seen as special cases of the same mechanism. We argue that imitation learning of context-dependent skills implies complex inferences to solve what we call the "motor Gavagai problem", which can be viewed as a generalization of the so-called "language Gavagai problem". In a full symbolic framework where percepts and actions are continuous, this allows us to articulate that language may be acquired out of generic sensorimotor imitation learning mechanisms primarily dedicated at solving this motor Gavagai problem. Through the use of a computational model, we illustrate how non-linguistic and linguistic skills can be learnt concurrently, seamlessly, and without the need for symbols. We also show that there is no need to actually represent the distinction between linguistic and non-linguistic tasks, which rather appears to be in the eye of the observer of the system. This computational model leverages advanced statistical methods for imitation learning, where closed-loop motor policies are learnt from human demonstrations of behaviours that are dynamical responses to a multimodal context. A novelty here is that the multimodal context, which defines what motor policy to achieve, includes, in addition to physical objects, a human interactant which can produce acoustic waves (speech) or hand gestures (sign language). A book chapter was written and published [66] and a journal article was submitted.

6.3.5. COSMO ("Communicating about Objects using Sensory-Motor Operations"): a Bayesian modeling framework for studying speech communication and the emergence of phonological systems

Participants: Clément Moulin-Frier, Jean-Luc Schwartz, Julien Diard, Pierre Bessière.

This work began with the PhD thesis of Clement Moulin-Frier at GIPSA-Lab, Grenoble, France, supervised by Jean-Luc Schwartz (GIPSA-Lab, CNRS), Julien Diard (LPNC, CNRS) and Pierre Bessière (College de France, CNRS). A few papers were finalized during his post-doc at FLOWERS in 2012. Firstly, an international journal paper based on the PhD thesis work of Raphael Laurent (GIPSA-Lab), extending Moulin-Frier's model, was published [25], and a commentary in *Behavioral and Brain Sciences* was accepted but not yet published [68]. Both these papers provide computational arguments based on a sensory-motor cognitive model to feed the age-old debate of motor vs. auditory theories of speech perception. Secondly, in another journal paper under the submission process, we attempt to derive some properties of phonological systems (the sound systems of human languages) from the mere properties of speech communication. We introduce a model of the cognitive architecture of a communicating agent, called COSMO (for "Communicating about Objects using Sensory-Motor Operations") that allows expressing in a probabilistic way the main theoretical trends found in the speech production and perception literature. This allows a computational comparison of these theoretical trends, helping to identify the conditions that favor the emergence of linguistic codes. We present realistic simulations of phonological system emergence showing that COSMO is able to predict the main regularities in vowel, stop consonant and syllable systems in human languages.

6.3.6. Recognizing speech in a novel accent: the Motor Theory of Speech Perception reframed

Participants: Clément Moulin-Frier, Michael Arbib.

Clément Moulin-Frier engaged this work with Michael Arbib during his 6-month visit in 2009 at the USC Brain Project, University of Southern California, Los Angeles, USA. An international journal paper is still under the revision process, in which we offer a novel computational model of foreign-accented speech adaptation, together with a thorough analysis of its implications with respect to the motor theory of speech perception.

6.3.7. Learning Simultaneously New Tasks and Feedback Models in Socially Guided Robot Learning

Participants: Manuel Lopes, Jonathan Grizou, Thomas Cederborg, Pierre-Yves Oudeyer.

We have developed a system that allows a robot to learn simultaneously new tasks and feedback models from ambiguous feedback in the context of robot learning by imitation. We have considered an inverse reinforcement learner that receives feedback from a user with an unknown and noisy protocol. The system needs to estimate simultaneously what the task is, and how the user is providing the feedback. We have further explored the problem of ambiguous protocols by considering that the words used by the teacher have an unknown relation with the action and meaning expected by the robot. This allows the system to start with a set of known symbols and learn the meaning of new ones. We have conducted human-robot interaction experiments where the user teaches a robot new tasks using natural speech with words unknown to the robot. The robot needs to estimate simultaneously what the task is and the associated meaning of words pronounced by the user. We have computational results showing that: a) it is possible to learn the task under unknown and noisy feedback, b) it is possible to reuse the acquired knowledge for learning new tasks and c) even in the presence of a known feedback, the use of extra unknown feedback signals while learning improves learning efficiency and robustness to mistakes. This algorithm has been applied on discrete and continuous problems and tested in a real world experiment using spoken words as feedback signals. A article to be submitted to a journal is currently being written.

6.3.8. Active Learning for Teaching a Robot Grounded Relational Symbols

Participants: Johannes Kulick, Tobias Lang, Marc Toussaint, Manuel Lopes.

The present work investigates an interactive teaching scenario, where a human aims to teach the robot symbols that abstract geometric (relational) features of objects. There are multiple motivations for this scenario: First, state-of-the-art methods for relational Reinforcement Learning demonstrated that we can successfully learn abstracting and well-generalizing probabilistic relational models and use them for goal-directed object manipulation. However, these methods rely on given grounded action and state symbols and raise the classical question Where do the symbols come from? Second, existing research on learning from human-robot interaction has focused mostly on the motion level (e.g., imitation learning). However, if the goal of teaching is to enable the robot to autonomously solve sequential manipulation tasks in a goal-directed manner, the human should have the possibility to teach the relevant abstractions to describe the task and let the robot eventually leverage powerful relational RL methods (see Figure 29). We formalize human-robot teaching of grounded symbols as an Active Learning problem, where the robot actively generates geometric situations that maximize his information gain about the symbol to be learnt. We demonstrate that the learned symbols can be used in a relational RL framework for the robot to learn probabilistic relational rules and use them to solve object manipulation tasks in a goal-directed manner. [44].

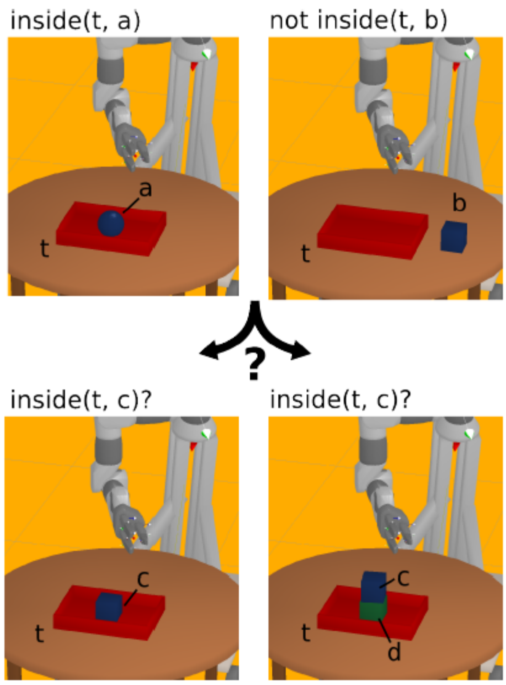


Figure 29. Active learning of symbol descriptions on a real world robot.

6.3.9. *Multimodal Conversational Interaction with a Humanoid Robot*

Participants: Adam Csapo, Emer Gilmartin, Jonathan Grizou, JingGuang Han, Raveesh Meena, Dimitra Anastasiou, Kristiina Jokinen, Graham Wilcock.

The paper presents a multimodal conversational interaction system for the Nao humanoid robot. The system was developed at the 8th International Summer Workshop on Multi-modal Interfaces, Metz, 2012. We implemented WikiTalk, an existing spoken dialog system for open-domain conversations, on Nao. This greatly extended the robot's interaction capabilities by enabling Nao to talk about an unlimited range of topics. In addition to speech interaction, we developed a wide range of multimodal interactive behaviours by the robot, including face- tracking, nodding, communicative gesturing, proximity detection and tactile interrupts. We made video recordings of user interactions and used questionnaires to evaluate the system. We further extended the robot's capabilities by linking Nao with Kinect. This work was presented in [34].

6.4. Other applications

6.4.1. *Real-time Reaction-Diffusion Simulation: a Machine Learning Technique*

Participants: Thomas Degris, Nejib Zemzemi.

Carmen is an Inria team working on modeling the electrical activity of the human heart. Their models are mainly based on reaction-diffusion equations. These methods are expansive in terms of computational costs which limits their use in practice. More specifically, some recent chirurgical intervention techniques on the heart (atrial ablation) requires to identify the source of the electrical wave. Finding such sources requires an optimization procedure. Using classical methods, this procedure is very heavy computationally.

In this project, our goal is to reduce the computational cost using supervised learning techniques. The idea is to replace the incremental resolution of partial differential equations by more suitable data structures for real-time running. Starting from data generated by simulating different excitations scenari on a human atria, this data is afterwords used as a training data set for machine learning algorithms. This approach will allow a faster optimization procedure.

This work is in collaboration with Nejib Zemzemi from the Inria Carmen team. This project is in preliminary steps.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

7.1.1. *Fondation Cartier pour l'Art Contemporain*

The team has been collaborating with Fondation Cartier pour l'Art Contemporain in the context of the elaboration of the exhibition "Mathematical: A Beautiful Elsewhere" (http://fondation.cartier.com), to be held from October 2011 to March 2012, as well as with artist David Lynch, to build the robotic installation/experiment Ergo-Robots/FLOWERS Fields. This robotic installation illustrates, as well as allows to experiment in a realistic setup on the long term, computational models of curiosity-driven learning, human-robot interaction and language formation. Fondation Cartier participated to the funding of this experiment/installation. A dedicated web page is available at: http://flowers.inria.fr/ergo-robots.php

7.1.2. Honda Research Institute USA Inc.

Alexander Gepperth is collaborating with Honda Research Institute USA Inc. to implement and evaluate a real-time pedestrian detection and pose classification system with the goal of creating an industrial product in the coming years. Particular aspects of the project are robustness and real-time capability. Robustness is approached by the use of state-of-the-art image feature representations, a sophisticated hierarchy of linear and non-linear support vector classifiers, and dedicated tracking algorithms. Real-time capability is ensured by running the time-critical parts of the whole-image search on a GPU. A particular focus of the project is the use of synthetically rendered pedestrian images for detector training, which ameliorates the problem of insufficient training data. This work has been submitted to the "International Conference On Computer Vision and Pattern Recognition" (CVPR) as well as the "Intelligent Vehicles Symposium" (IV). Honda Research Institute USA Inc. support Alexander Gepperth by financing a post-doctoral researcher at ENSTA ParisTech during one year, grant volume: 50.000USD.

7.1.3. Robert Kostal GmbH

Alexander Gepperth has collaborated with Robert Kostal GmbH, Dortmund (Germany) on the subject of real-time pose recognition from 3D camera data. This project was conducted mainly through an internship student financed by Robert Kostal GmbH.

7.1.4. Honda Research Institute Europe GmbH

Alexander Gepperth and Louis-Charles Caron have collaborated with Honda Research Institute Europe GmbH, Offenbach (Germany) on the subject of real-time shape recognition for robotics. This project was conducted through an internship student financed by Honda Research Institute Europe GmbH, and through the visit of Louis-Charles Caron to Honda Research Institute Europe GmbH in summer 2012.

7.1.5. Pal Robotics

Freek Stulp is continuing his collaboration with Pal Robotics in Barcelona to implement and evaluate the use of Dynamic Motion Primitives on the commercial mobile platform 'REEM'. A particular focus of this project is to compare the respective advantages of motion primitives and sampling-based motion planning approaches in the context of human-robot interaction. Pal Robotics is supporting Freek Stulp by co-financing travel costs for regular project meetings in Barcelona: http://www.pal-robotics.com/blog/freek-stulp-visited-pal-robotics/. In 2012 this collaboration has lead to a paper at Humanoids [45], and a video at IROS, which was selected for an interactive session, “in consideration of the quality of your work”.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. ADT CARRoMan

The ADT project CARRoMan started in november 2012 (recruitement of Antoine Hoarau). Autonomous human-centered robots, for instance robots that assist people with disabilities, must be able to physically manipulate their environment. There is therefore a strong interest within the FLOWERS team to apply the developmental approach to robotics in particular to the acquisition of sophisticated skills for manipulation and perception. ENSTA-ParisTech has recently acquired a Meka humanoid robot dedicated to human-robot interaction, and which is perfectly fitted to this research. The goal of this ADT is to install state-of-the-art software architecture and libraries for perception and control on the Meka robot, so that this robot can be jointly used by FLOWERS and ENSTA. In particular, we want to provide the robot with an initial set of manipulation skills. The engineer will develop a set of demos, which demonstrate the capabilities of the Meka, and provide a basis on which researchers can base their experiments.

8.1.2. CRA ARAUI

A Conseil Régional d'Aquitaine Project (ARAUI, 2011-) began, coordinated by Manuel Lopes entitled *Apprentissage Automatique en Robotique pour l'Adaptation aux Utilisateurs a Travers L'Interaction*. It will fund 50% of a 3 years PhD student and funding of 5500 euros for equipment.
The objective of ARAUI is the creation of robots that initiate autonomously the execution of frequent tasks after learning about the user's preferences through repeated interactions. Particularly these robots will act as personal companions or helpers and will be able to perform shared tasks with humans.

The long-term view of this project is that of a robot that comes out of the box with general purpose motor and sensory skills and then is adapted to each user's preferences and needs to achieve autonomous behavior. The major challenge is how to equip machines with such adaptability and learning capabilities. Until now machines are programmed by skilled engineers to perform a specific task and learning new tasks is not possible. Even in a restricted industrial setting the need for robots that can be more easily re-programmed to new tasks and environments has lead to research programs on flexible manufacturing that consider frequent changes in tasks and close (physical) interactions with human operators.

8.1.3. CRA ACROBATE

The Conseil Régional d'Aquitaine Project (ACROBATE, 2009-) continued, involving Thomas Cederborg and Pierre-Yves Oudeyer. The funding contributes with 50% funding for a 3 years PhD student. The objective of ACROBATE is to study mechanisms and models that can allow a robot to learn in a unified manner context-dependant motor skills and linguistic skills through interactions with humans.

8.1.4. ADT Acrodev

The ADT project (Acrodev, 2010-) continued, involving Paul Fudal, Haylee Fogg, Olivier Ly and Pierre-Yves Oudeyer. The Inria ADT funds two engineers for two years. The objective of Acrodev is on the one hand to build up re-usable software architectures for embedded control of Acroban-like robots, and on the other hand to explore novel morphologies in particular for the feet, hands and head of Acroban-like robots.

8.1.5. Collaboration with Labri/Unvi. Bordeaux I

We continued to collaborate with the Rhoban group at Labri/CNRS/Univ. Bordeaux I, and in particular Olivier Ly and Hugo Gimbert, about the design of bio-inspired compliant robotic morphologies, such as around the Acroban humanoid robot. The goal is to study both how properties of the body can facilitate motor control, and how to experiment and design such bodies with rapid prototyping methods.

8.1.6. Collaboration with Labri/Univ. Bordeaux I and Institut de Neurosciences Cognitives et Integratives d'Aquitaine

The collaboration with Olivier Ly, from Labri and Univ. Bordeaux I, as well as with Jean-René Cazalets, Christophe Halgand and Etienne Guillaud from Institut de Neurosciences Cognitives et Integratives d'Aquitaine, Bordeaux continued. The goal is to compare properties of the postural balance, and its relation to morphology and distributed control, in humans and in the humanoid Acroban (developped in collaboration with Labri), which vertebral column and postural control shares several fundamental features with the human vertebral column, and using the "Plateforme d'analyse de la motricité" available at the Institut de Neurosciences Cognitives et Integratives d'Aquitaine. This collaboration involves Matthieu Lapeyre and Pierre-Yves Oudeyer.

8.2. National Initiatives

8.2.1. ANR MACSi

An ANR Project (MACSi, ANR Blanc 0216 02), coordinated by ISIR/Univesity Paris VI (Olivier Sigaud), on developmental robotics (motor learning, visual learning, and exploration algorithms on the ICub robot) continued. The MACSi project is a developmental robotics project based on the iCub humanoid robot and the Urbi open source software platform. It is funded an as ANR Blanc project from 2010 to 2012. The project addresses four fundamental challenges, led by four partners:

- How can a robot learn efficient perceptual representations of its body and of external objects given initially only low-level perceptual capabilities? Challenge leader : Inria-ENSTA-ParisTech FLOWERS (Paris).
- How can a robot learn motor representations and use them to build basic affordant reaching and manipulation skills? Challenge leader : ISIR-UPMC-Paris 6 (Paris). ISIR hosts the iCub humanoid robot on which the achievements will be evaluated.
- What guidance heuristics should be used to explore vast sensorimotor spaces in unknown changing bodies and environments? Challenge leader : Inria-ENSTA-ParisTech FLOWERS (Bordeaux).
- How can mechanisms for building efficient representations/abstractions, mechanisms for learning manipulation skills, and guidance mechanisms be integrated in the same experimental robotic architecture and reused for different robots? Challenge leader : GOSTAI company (Paris).

Web site: http://macsi.isir.upmc.fr/

8.2.2. Quasimetric approach to probabilistic optimal control (LPPA)

- Jean-Luc Schwartz[1], Julien Diard[2], Pierre Bessire[3], Raphael Laurent[4], **1: GIPSA-Lab, Grenoble University, CNRS. 2: LPNC, Grenoble University, CNRS. 3: LPPA, Collège de France, CNRS. 4: GIPSA-Lab, Grenoble University.** Clément Moulin-Frier is continuing his collaborative work with people he worked with during his PhD thesis at GIPSA-Lab. See the section entitled ”COSMO (”Communicating about Objects using Sensory-Motor Operations”): a Bayesian modeling framework for studying speech communication and the emergence of phonological systems” for more information.
- Jacques Droulez, Steve N'Guyen, **Laboratoire de Physiologie de Perception et de l'Action (LPPA), College de France, Paris**. Clément Moulin-Frier is continuing his collaborative work with people he worked with during his post-doc in 2011 at LPPA, College de France. See the section entitled ”Probabilistic optimal control: a quasimetric approach” for more information.

8.2.3. Collaboration and technological transfer with Laboratoire de Physiologie de la Perception et de l'Action (LPPA)

A collaboration is in progress with Jacques Droulez and Steve Nguyen from Laboratoire de Physiologie de la Perception et de l'Action (LPPA), Paris. Poppy represents for them a humanoid platform very interesting because it is relatively flexible and versatile, with more similar proportions to that of humans, which facilitate comparison with the experimental results obtained in humans. The laboratory will evaluate this platform probabilistic methods of control of balance and locomotion.

In the short term the first experimental project with Poppy will test methods of management support, in the case of restoration of balance, in the case of walking to correct or prepare a change of direction. This project will be initiated in the framework of a long internship of master 2 that starts in January. In the future, we would also like to evaluate motor controllers compliant, and learning algorithms. This collaboration involves Matthieu Lapeyre and Pierre-Yves Oudeyer.

8.3. European Initiatives

8.3.1. FP7 Projects

8.3.1.1. EXPLORERS

Program: ERC Starting Grant

Project acronym: EXPLORERS

Project title: Exploring Epigenetic Robotics: Raising Intelligence in Machines

Duration: 12/2009-11/2014

Coordinator: Pierre-Yves Oudeyer

Abstract: In spite of considerable and impressive work in artificial intelligence, machine learning, and pattern recognition in the past 50 years, we have no machine capable of adapting to the physical and social environment with the flexibility, robustness and versatility of a 6-months old human child. Instead of trying to simulate directly the adult's intelligence, EXPLORERS proposes to focus on the developmental processes that give rise to intelligence in infants by re-implementing them in machines. Framed in the developmental/epigenetic robotics research agenda, and grounded in research in human developmental psychology, its main target is to build robotic machines capable of autonomously learning and re-using a variety of skills and know-how that were not specified at design time, and with initially limited knowledge of the body and of the environment in which it will operate. This implies several fundamental issues: How can a robot discover its body and its relationships with the physical and social environment? How can it learn new skills without the intervention of an engineer? What internal motivations shall guide its exploration of vast spaces of skills? Can it learn through natural social interactions with humans? How to represent the learnt skills and how can they be re-used? EXPLORERS attacks directly those questions by proposing a series of scientific and technological advances: 1) we will formalize and implement sophisticated systems of intrinsic motivation, responsible of organized spontaneous exploration in humans, for the regulation of the growth of complexity of learning situations; 2) intrinsic motivation systems will be used to drive the learning of forward/anticipative sensorimotor models in high-dimensional multimodal spaces, as well as the building of reusable behavioural macros; 3) intrinsically motivated exploration will be coupled with social guidance from non-engineer humans; 4) an information-theoretic framework will complement intrinsically motivated exploration to allow for the inference of body maps; 5) we will show how learnt basic sensorimotor skills can be re-used to learn the meaning of early concrete words, pushing forward human-robot mutual understanding. Furthermore, we will setup large scale experiments, in order to show how these advances can allow a high-dimensional multimodal robot to learn collections of skills continuously in a weeks-to-months time scale. This project not only addresses fundamental scientific questions, but also relates to important societal issues: personal home robots are bound to become part of everyday life in the 21st century, in particular as helpful social companions in an aging society. EXPLORERS' objectives converge to the challenges implied by this vision: robots will have to be able to adapt and learn new skills in the unknown homes of users who are not engineers. The ERC EXPLORERS is a central scientific driver of the FLOWERS team.

8.4. International Initiatives

8.4.1. Inria International Partners

- Luis Montesano, **University of Zaragoza, Spain**. Manuel Lopes collaborated with Luis Montesano on several topics. Recently on active learning approaches for grasping point learning [103] and clustering activities.
- Francisco Melo **Instituto Superior Técnico, Portugal**. Manuel Lopes collaborated with Francisco Melo on the development of active learning for inverse reinforcement learning. Recent developments consider the extension to more cues available to the learner and sampling complexity on the algorithm.
- José Santos-Victor, **Instituto Superior Técnico, Portugal**. Manuel Lopes collaborated with José Santos-Victor on the extension of affordances models to higher levels of representations, e.g. relational models.
- Maya Cakmak, Andrea Thomaz, **Georgia Tech, USA**. Manuel Lopes collaborated with Maya Cakmak on the development of optimal teaching algorithms for sequential decision problems (modeled as markov decision processes). The algorithm provides optimal demonstrations for systems that learn using inverse reinforcement learning. The joint work considers not only the algorithmic aspects but also a comparison with human behavior and the possibility of using insights from the algorithm to elicit better teaching behavior on humans [32].

- Marc Toussaint, Tobias Lang, **Free University of Berlin, Germany**. Manuel Lopes and Pierre-Yves Oudeyer are collaborating with FUB in the unification of exploration algorithms based on intrinsic motivation with methods for exploration in reinforcement learning such as R_{max}. We intend to develop a general framework for exploration in non-stationary domains [46]. Another project consider how to learn efficient representation for robotic hierarchical planning [44].
- Todd Hester and Peter Stone, **University of Texas, USA** (2012 -)
Peter Stone is a leading expert on reinforcement learning applied to real robots (he won the RobotCup competition several times) and to multi-agent problems. We started this collaboration by introducing a new method to automatically select the best exploration strategy to use in a particular problem [42]. Future directions of the collaboration will include ad-hoc teams, exploration in continuous space and human-guided machine learning.
- Jacqueline Gottlieb and Adrien Baranes, **Columbia University, New-York, US**. Pierre-Yves Oudeyer and Manuel Lopes continued a collaboration with Jacqueline Gottlied, neuroscientist at Columbia University and specialist of visual attention and exploration in monkeys, and Adrien Baranes, postdoc in Gottlieb's lab and previously working in Flowers team. An experimental set-up with brain imaging and behavioural observations of monkeys, and made to evaluate new families of computational models of visual attention and exploration (some of which developped in the team around the concept of intrinsic motivation) is being elaborated.
- Louis ten Bosch, **Radboud University, The Netherlands**. Pierre-Yves Oudeyer and David Filliat continued to work with Louis ten Bosch on the modelling of multimodal language acquisition using techniques based on Non-Negative Matrix Factorization. We showed that these techniques can allow a robot to discover audio-video invariants starting from a continuous unlabelled and unsegmented flow of low-level auditory and visual stimuli. A journal article is in preparation.
- Britta Wrede, Katharina Rohlfing, Jochen Steil and Sebastian Wrede, **Bielefeld University, Germany**, Jun Tani **KAIST, South Korea**. Pierre-Yves Oudeyer collaborated with Wrede, Rohlfing, Steil, Wrede and Tani on the elaboration of a novel conceptual vision of teleoogical language and action development in robots. This led to the publication of a joint workshop article [64].
- Michael A. Arbib, **University of Southern California (Los Angeles, USA)**. Clément Moulin-Frier is continuing his collaborative work with Michael Arbib since his 6-month visit at USC in 2009. See the section entitled “Recognizing speech in a novel accent: the Motor Theory of Speech Perception reframed” for more information.
- Paul Vogt (**Tillburg University, The Netherlands**), Linda Smith (**Indiana University, Bloomington, US**), Aslo Ozyurek (**Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands**), Tony Belpaeme (**University of Plymouth, UK**). Pierre-Yves Oudeyer began collaboration with partners of the NWO SCMSC project to set up a research network on modeling of social cognition and symbolic communication.
- Michael Gienger from **Honda Research Institute Europe**. Alexander Gepperth collaborated with Principal Scientist Dr.Michael Gienger from Honda Research Institute Europe GmbH about robotic grasping: this activity will result in a jointly supervised internship ("stage de fine d'études") and a publication.
- Ursula Korner from **Honda Research Institute Europe**. Alexander Gepperth collaborated with Dr. Usula Körner of Honda Research Institute Europe GmbH, Offenbach (Germany), on the topic of biologically inspired learning architectures for visual categorization of behaviorally relevant entities. This work is intended to be summitted to the International Conference on Development and Learning, as well as the journal "Frontiers in Cognitive Systems".
- Michael Garcia Ortiz, **Laboratory for Cognitive Robotics (CoR-Lab) in Bielefeld, Germany**. Alexander Gepperth collaborated with Michael Garcia Ortiz, a PhD student from the Laboratory for Cognitive Robotics (CoR-Lab) in Bielefeld, Germany, on the exploitation of scene context for object detection in intelligent vehicles. This collaboration resulted in the submission of a journal publication to the journal ”Neurocomputing”.

- Martha White and Richard Sutton from the **University of Alberta, Canada**. Thomas Degris collaborated with Martha White and Richard Sutton on the paper "Off-Policy Actor–Critic" [38].
- Patrick Pilarski and Richard Sutton from the University of Alberta (Canada). Thomas Degris collaborated with Patrick Pilarski on the following papers: "Model-Free Reinforcement Learning with Continuous Action in Practice" [37], "Apprentissage par Renforcement sans Modèle et avec Action Continue" [65], "Dynamic Switching and Real-time Machine Learning for Improved Human Control of Assistive Biomedical Robots" [57], "Towards Prediction-Based Prosthetic Control" [58], and "Prediction and Anticipation for Adaptive Artificial Limbs" [27].
- Joseph Modayil from the **University of Alberta, Canada**. Thomas Degris collaborated with Joseph Modayil on the following paper: "Scaling-up Knowledge for a Cognizant Robot" [35].
- Ashique Rupam Mahmood from the **University of Alberta, Canada**. Thomas Degris collaborated with Ashique Rupam Mahmood on the following paper: "Tuning-Free Step-Size Adaptation" [50].

8.5. International Research Visitors

8.5.1. Visits of International Scientists

- Andrew Barto, Reinforcement learning and intrinsic motivation, University of Massachusetts Amherst, USA (oct 2012)
- Adam White, Reinforcement Learning and Artificial Intelligent group, Computing Science department of the University of Alberta, Canada (September 2012)
- Joseph Modayil, Reinforcement Learning and Artificial Intelligent group, Computing Science department of the University of Alberta, Canada (September 2012)
- Akihiko Yamaguchi, Robotics Lab of Prof. Ogasawara at NAIST in Japan (march 2012)
- Todd Hester, RL and Robotics Lab, Univ. Texas, US (may, june, july 2012)
- Louis ten Bosh, Speech processing, Univ. Radboud, The Netherlands (june 2012)
- Robert Saunders, Design Lab, Faculty of Architecture, University of Sydney, Australia (september 2012)
- Adrien Baranes, Columbia University, NY, USA (october 2012)
- Joshka Boedecker, Asada Lab, Osaka University, Japan (october 2012)
- Olivier Georgeon, Univ. Lyon, France (november 2012)

8.5.2. Internships

- Gennaro Raiola, MSc. Student from Università degli Studi di Napoli Federico II. Parameterized skills are able to map parameters of the task (for instance the 2D position of an object on a table) to the appropriate parameters of a policy for achieving this task. In this project, we use imitation learning to train a Dynamic Movement Primitive (DMP) with several observed trajectories. To achieve generalization, the basis functions in the DMP are expanded so that they span the space of the task relevant parameters. The resulting algorithm is applied to human reaching data, and to generalizing skills on the Nao robot.
- Laura Vogelaar, visiting student from GeorgiaTech and Carnegie Mellon University. Within a stochastic optimization context, we use clustering algorithms to determine features that are relevant to minimizing the cost of executing a skill. Our objective is to enable a robot to autonomously expand its libraries of skills, whilst simultaneously learning which skills can be successfully executed in which contexts.

8.5.3. Visits to International Teams

- Manuel Lopes (December 2012), Willow Garage, Palo Alto, USA: visit to Maya Cakmak to discuss tutoring systems and human-robot internaction.

- Manuel Lopes (December 2012), Bosch Research, Palo Alto, USA: visit to Dejan Pangercic to discuss active learning and human-robot interaction.
- Manuel Lopes (December 2012), Berkely University, USA: visit to Pieter Abbeel to discuss safe exploration methods and inverse reinforcement learning.
- Manuel Lopes (December 2012),
- Clément Moulin-Frier (November 2012), UC Merced, USA: visit to Anne Warlaumont's lab at UC Merced, to discuss about the role and the computational modeling of infraphonology in infant language development. The aim is to initiate a collaboration with Anne Warlaumont and D. Kimbrough Oller (University of Memphis, USA) to computationally study the possible role of intrinsic motivations in infraphonological exploration.
- Olivier Mangin (17/10/2012), Instituto Superior Técnico, Lisbone, Portugal
- Thomas Degris (June 2012), Reinforcement Learning and Artificial Intelligent group, Computing Science department of the University of Alberta, Canada (June 2012)

8.5.4. Participation to Summer/Winter School

- Jonathan Grizou participated to e'NTERFACE 2012, July, 2nd - July, 27th 2012, SUPELEC, Metz, France The 8th International Summer Workshop on Multimodal Interfaces took place on the SUPELEC campus of Metz, France. This one month summer school brought together more than 70 students and experts to work together and foster the development of tomorrow's multimodal research community. Jonathan Grizou enrolled in the Project P1 : "Speech, gaze and gesturing – multimodal conversational interaction with Nao robot", supervised by Graham Wilcock and Kristiina Jokinen (University of Helsinki). This summer school lead to a join publication by the members of the project P1 at the CogInfoCom 2012 conference [34].
- Jonathan Grizou and Fabien Bénureau participated to the IM-CLeVeR/FIAS Winter School on "Intrinsic Motivation: From Brains to Robots", December 3-8, 2012, Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany. The school brought together 25 students in the field of intrinsic motivation as well as leaders in the field (among which, Andrew Barto, Minoru Asada, Peter Redgrave, Giorgio Metta and others). Students' time was divided between keynotes in the morning and project work in the afternoon, supervised by the speakers and the school organizers. The school was an opportunity to meet and discuss with researchers and PhD students. It also allowed us to explain and disseminate our work; Pierre-Yves Oudeyer, notably, was an invited speaker. Jonathan Grizou took part in the project "Intrinsic Motivation in Active Perception" while Fabien Benureau participated in "Playful Acquisition of Basic Behavioral skills Machine". The results of the school are highly positive, and some scientific collaborations may directly stem from this event in the future.

9. Dissemination

9.1. Animation of the Scientific Community

9.1.1. Editorial boards

Pierre-Yves Oudeyer has worked as Editor of the IEEE CIS AMD Newsletter, and member of the IEEE CIS Technical Committee on Autonomous Mental Development.

Pierre-Yves Oudeyer has worked as Associate Editor for IEEE Transactions on Autonomous Mental Development, Frontiers in Neurorobotics (Frontiers Foundation), International Journal of Social Robotics (Springer).

Pierre-Yves Oudeyer has worked as member of the editorial board of the book series Advances in Interaction Studies, John Benjamins Publishing Company.

9.1.2. Steering committees

Pierre-Yves Oudeyer has worked as member of the Steering Committee of the International Conference on Epigenetic Robotics, and participated to the setting up of the second joint conference with IEEE ICDL, i.e. the IEEE ICDL/Epirob conference that was held in San Diego, US.

Manuel Lopes participated in the steering committee of the IEEE TC on Robot Learning.

9.1.3. Conference Organization

Pierre-Yves Oudeyer co-organized the third International Workshop on Human Behavior Understanding, and was co-editor of the proceedings [67], [59]: http://www.cmpe.boun.edu.tr/hbu/2012/ (together with A. A. Salah, Cetin Mericli and Javier Ruiz-del-Solar).

9.1.4. Program Committees

Freek Stulp was on the program committee of the IEEE International Conference on Development and Learning/Epigenetic Robotics.
Pierre-Yves Oudeyer was a member of the following program committees: IEEE ICDL-EPIROB 2012; IEEE RAS International Conference on Humanoid Robots (HUMANOIDS); 3rd International Workshop on Human-Behaviour Understanding (HBU).
Manuel Lopes was on the program committee of STAIRS/ECAI, AAAI, Inter. Conf. on Autonomous Robot Systems and Competitions and European Workshop on Reinforcement Learning (EWRL).

9.1.5. Journal Reviews

David Filliat reviewed papers for the journals: Autonomous Robots, Robotics and Autonomous Systems and Journal of Visual Communication and Image Representation.
Freek Stulp reviewed papers for the journals: IEEE Transactions on Robotics, Transactions on Mechatronics, IEEE Transactions on Control Systems Technology.
Pierre-Yves Oudeyer reviewed papers for the journals:IEEE Transactions on Autonomous Mental Development, Neural Networks.
Manuel Lopes reviewed for the journals: IEEE Transactions on Robotics, IEEE Transactions on Autonomous Mental Development, Robotics and Autonomous Systems and Advanced Robotics.
Alexander Gepperth was a reviewer for the journals "Cognitive Computation", "Neural Processing Letters" and "Intelligent Transportation Systems".
Thomas Degris was a reviewer for the Neural Computation Journal and Revue d'Intelligence Artificielle.

9.1.6. Conference Reviews

Alexander Gepperth reviews articles for the International Conference on Neurally Inspired Processing (ICONIP) and the ACM/IEEE HRI conference.
Freek Stulp was reviewer for IEEE International Conference on Robotics and Automation (ICRA), IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE International Conference on Development and Learning (ICDL-EpiRob), IEEE-RAS International Conference on Humanoid Robots (Humanoids), IEEE International Symposium on Robot and Human Interactive Communication (ROMAN).
David Filliat was reviewer for IEEE International Conference on Robotics and Automation (ICRA), IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Advanced Concepts for Intelligent Vision Systems (ACIVS), International Conference on Control, Automation, Robotics and Vision (ICARCV) and the 10th International IFAC Symposium on Robot Control (SYROCO).
Pierre-Yves Oudeyer was a reviewer for the conferences IEEE ICDL-EPIROB, Humanoids, ICRA.
Pierre Rouanet has reviewed papers for the IEEE 2012 ICRA conference, IEEE 2011 IROS conference and IEEE HRI 2011 workshop "Expectations in intuitive HRI".
Thomas Cederborg has has reviewed one paper for the IROS 2012 conference, one paper for the ICDL-EpiRob 2012 conference and one paper for the HBU 2012 conference.
Jonathan Grizou has reviewed papers for the IEEE ICDL 2012 conference.
Sao Mai Nguyen has reviewed a paper for IROS 2012, ICDL 2012 and a workshop of IROS 2012.

Olivier Mangin has reviewed papers for the Humanoid 2012 conference, for the Human Behavior Understanding 2012 workshop, and for the ICRA 2013 conference.
Thomas Degris has reviewed papers for the ICDL Epirob 2012 conference.

9.1.7. PhD Jury

Pierre-Yves Oudeyer was rapporteur in the PhD jury of Matthias Rolf (Bielefeld University, Germany), for its PhD entitled “Goal babbling for an efficient bootstrapping of inverse models in high dimensions”, as well as rapporteur in the PhD jury of Duong Dang (LAAS CNRS, France), for its PhD entitled “Manipulation et locomotion en robotique humanoïde avec optimisation en temps réel des pas”, and also participated to the PhD jury of John Nassour (Univ. Versailles, France; and TUM, Germany), for its PhD entitled “Success-failure learning for humanoid: study on bipedal walking”.
David Filliat was rapporteur in the PhD jury of Mathieu Dubois (Méthodes probabilistes basées sur les mots visuels pour la reconnaissance de lieux par un robot mobile, 20/02/12), of Thomas Moulard (Optimisation numérique pour la robotique et exécution de trajectoires référencées capteurs, 17/09/12), of Fengchun Dong (Vision sensor design and evaluation for autonomous navigation, 22/11/12), of Ahmad Mohammed Hasasneh (Robot semantic place recognition based on deep belief networks and a direct use of tiny images, 23/11/12) and participated in the jury of Pierre Rouanet (Apprendre à un robot à reconnaître des objets visuels nouveaux et à les associer à des mots nouveaux : le rôle de l’interface, 04/03/12).
Manuel Lopes acted as member of the advising committee for the PhD thesis of: Pedro Sequeira entitled ”Biologically-inspired Mechanisms to Enhance Learning in Autonomous Agents”, Instituto Superior Técnico, Lisbon, Portugal, and Salomon Ramire entitled ”Active Vision in the Peripersonal Space for Humanoid Robots”, University of Plymouth, England.
Alexander Gepperth is/will be rapporteur in the PhD jury of Michael Garcia Ortiz, for his PhD entitled "Driver Behavior Prediction in intelligent vehicles", to be submitted to the university of Bielefeld, Germany in January 2013.

9.1.8. Expertise

David Filliat reviewed projects for the ’Programme Evaluation-orientation de la Coopération Scientifique (ECOS)’.
Pierre-Yves Oudeyer was expert for the European Commission for review and evaluations of several FP7 projects and calls (ICT and FET). He was also reviewer for ANR projects’, and was a member of Commission de Développement Technologique, Inria Bordeaux Sud-Ouest.
Manuel Lopes participated in the Cost-Gtri: Groupe de travail des relations internationales du comite d’orientation scientifique et technologique.
Thomas Degris reviewed a project proposal in Reinforcement Learning for the The Netherlands Organisation for Scientific Research (NWO).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

License: Introduction à Matlab, 21 heures. L3 - ENSTA ParisTech (Alexander Gepperth)

License: Traitement numérique du signal , 21 heures. L3 - ENSTA ParisTech (Alexander Gepperth)

Master : Robotique de Comnpagnie, 12 heurs. M2, ENSTA - Paris Tech (Manuel Lopes).

Master : Robotique Mobile, 24 heures. M2, ENSTA - ParisTech (David Filliat).

Master : Vision pour la robotique, 12 heures. M2, University Pierre et Marie Curie (David Filliat).

License : Introduction to Matlab, 21 heures. L3, ENSTA - ParisTech (David Filliat).

License : Projet informatique, 21 heures. L3, ENSTA - ParisTech (David Filliat).

Licence 2 : Graphe, Langage, Cryptologie, 21 heures. Ple Universitaire Francais de Ho Chi Minh Ville

Master : Option Robotique, Projet Robot Autonome, 32 heures. ENSEIRB, Bordeaux, France.

Licence : Mathématique, 64 heures, niveau (L2), Pôle Universitaire Français, Ho Chi Minh ville (Olivier Mangin)

Licence : Programmation système, 30 heures, niveau (L2), Pôle Universitaire Français, Ho Chi Minh ville (Olivier Mangin)

9.2.2. Supervision

PhD & HdR :

HdR: Pierre-Yves Oudeyer defended his HdR, entitled "Developmental constraints on the evolution and acquisition of sensorimotor and social skills", at University of Bordeaux, 18 May, 2011.

HdR : David Filliat, Navigation, Perception et Apprentissage pour la robotique, University Pierre et Marie Curie, 12 Juillet 2011.

PhD defended by Pierre Rouanet ([21]), "Apprendre à un robot à reconnaître des objets visuels nouveaux et à les associer à des mots nouveaux : le rôle de l'interface", at University Bordeaux I, march 2012 (superv. Pierre-Yves Oudeyer).

PhD in progress : Louis-Charles Caron, Developmental learning in multimodal sensory-motor loops, started january 2012 (superv. Alexander Gepperth).

PhD in progress : Natalia Lyubova, A developmental approach to perception for a humanoid robot, started nov 2010 (superv. David Filliat).

PhD in progress : Matthieu Lapeyre, Developmental constraints for biped humanoid walking, started oct. 2010 (superv. Pierre-Yves Oudeyer and Olivier Ly).

PhD in progress : Mai Nguyen, Bootstrapping Intrinsically Motivated Learning with Human Demonstration, started oct. 2010 (superv. Pierre-Yves Oudeyer).

PhD in progress : Fabien Bénureau, Cumulative, hierarchical and intrinsically motivated learning of robot skills, started oct. 2010 (superv. Pierre-Yves Oudeyer).

PhD in progress : Jonathan Grizou, Fluid simultaneous learning of task and feedback models, started oct. 2011 (superv. Manuel Lopes and Pierre-Yves Oudeyer).

PhD in progress : Olivier Mangin, Learning of sensorimotor primitives with Non-Negative Matrix Factorization, started oct. 2010 (superv. Pierre-Yves Oudeyer).

PhD in progress : Thomas Cederborg, A unified view of context-dependant skill learning and language acquisition, started oct. 2009 (superv. Pierre-Yves Oudeyer).

Master internships and others:

- J. Grizou supervised the 6 month master 2 internship of Mathieu Duteil Master thesis project, coming from "Université Pierre et Marie Curie", where he studied Intelligent Systems and Robotics. The project consisted in detecting if humans provide non-verbal feedback during interactions with robots. Mathieu did several user studies to record the necessary data using several human-machine interaction protocols, recording the commands provided by the person, the sounds and the facial expressions. With this data he tested several kernel based methods to allow classifying the relevant feedback from the human.
- J. Grizou supervised two middle school students. The purpose of such one week internship is to allow pre high-school students to discover what being a researcher means as well as discovering some simple technological setup.
- F. Benureay supervised one middle school student, Victor Melançon, 16 years old, for one week internship. He worked on a visualization and control of the motion of robotic arm using a interface created during the internship using Processing. The goal was to elaborate an aesthetically pleasing visualization of the robot motion that would engage non-expert users in an interaction with the robot. The project was also the opportunity to learn the basics of programming and robot control.
- M. Nguyen supervised the internship of Thomas Huet (ENS Paris) for a 2 months internship entitled : "Learning Methods for Robotic Models in a Fishing Experiment".
- P-Y. Oudeyer supervised the 3 months master 1 internship of Sébastien Forestier (ENS Cachan). The project consisted in experimenting the SAGG-RIAC architecture on the Ergo-Robots, and it was shown to allow successful learning of inverse kinematics.

9.3. Invited talks

Pierre-Yves Oudeyer:

- (5th december 2012) *Developmental mechanisms for life-long autonomous learning in robots and humans*, FIAS Winter School on Intrinsic Motivation: From Brains to Robots, Frankfurt, Germany.
- (18th october 2012) Le rôle du corps en robotique développementale, Conférence "Corps et Robots", ARCO/IPac, Inria Nancy, France.
- (12th october 2012) *A Robotic Platform for Scalable Life-Long Learning Experiments*, IROS 2012 Workshop on Learning and Interaction in Haptic Robots, Vilamoura, Algarve, Portugal.
- (5th october 2012) *Developmental mechanisms for life-long autonomous learning*, INNS Symposium on Autonomous Learning, 2012 International Neural Network Society Winter Conference (INNS-WC2012).
- (6th september 2012) *The challenges of active exploration and learning in high-dimensional continuous spaces*, GdR CNRS Robotique et Neurosciences.
- (31st august 2012) *Developmental mechanisms for life-long autonomous learning in robots*, Frontiers of AI track, 20th European Conference on Artificial Intelligence (ECAI 2012).
- (6th july 2012) *Bootstrapping language development out of multimodal sub-symbolic sensorimotor learning in robots*, Symposium on Origins of Communication and Language, Epigenetic Modeling and Ethodological Observation, Konrad Lorenz Institute, Altenberg, Austria.
- (5th april 2012) *Developmental autonomous learning*, Séminaire du laboratoire L3i de l'Université de La Rochelle.
- (30th january 2012) *Les modèles robotiques: un nouveau langage pour comprendre le vivant*, Colloque "Mathématiques pour tous?", organisé par l'UNESCO en partenariat avec l'IHES et la Fondation Cartier pour l'Art Contemporain, UNESCO, Paris, France.
- (26th january 2012) *From the language Gavavai problem to the motor Gavagai problem: Modeling language acquisition as an instance of general multimodal context-dependant learning by imitation*, Workshop on Socio-Cognitive Mechanisms of Symbolic Communication, Tilburg University, Tilburg, The Netherlands.

Freek Stulp:

- (february 2012) Invited talk at a meeting of the EU IP project "HANDLE: Developmental pathway towards autonomy and dexterity in robot in-hand manipulation". Benicassim, Spain. Title: *Motion Primitives and Direct Reinforcement Learning for Robot Manipulation*.
- (november 2012) Invited talk at "Journée Evolution Artificielle Thématique" in Paris, France. Title: *From Episodic Reinforcement Learning and Quantum Mechanics to Evolutionary Optimization*'.

Manuel Lopes:

- (december 2012), *Active and Social Learning for Robots*, Bosch Research, Palo Alto, USA.
- (september 2012) *Autonomous Exploration Through Curiosity and Social Guidance*, Evo-Devo-Robo: Evolutionary Robotics and Developmental Robotics, GECCO, Philadelphia, USA.
- (july 2012) *Interactive Learning in Social Robots*, German-French Workshop: Perspectives on Cognitive Interaction and Technology, Bielefeld, Germany, 2012.
- (april 2012) *"Ces robots qui nous imitent"*, Unithé ou café, Bordeaux.

Thomas Degris:

- (march 2012) *Off-policy Actor-critic: Algorithm and empirical evaluations*, 7th Workshop on Reinforcement Learning, Policy Approximation, Barbados.
- (march 2012) *Formalizing Curiosity, Social interaction and Maturation*, 7th Workshop on Reinforcement Learning, Policy Approximation, Barbados.

9.4. Popularization

9.4.1. Popular Science Articles

Filliat, D. (2012) Vers une cartographie sémantique d'environnements intérieurs. Réalités Industrielles, Février 2012.
Oudeyer, P-Y. (2012) GX-29 n'est pas un objet comme les autres, Sciences et Avenir Hors-Série, dec/jan 2011, "Qu'est-ce-que l'homme". http://flowers.inria.fr/documents/SciencesEtAvenirDec2011.pdf
Oudeyer, P-Y. (2012) Les Robots Curieux, DocSciences 14, Alan Turing: La pensée informatique. http://www.pyoudeyer.com/DocSciencesErgoRobots12.pdf.

Ly, O., Oudeyer, P-Y., Langlois, A. (2012) Le déséquilibre de l'apprentissage, Interstices. http://interstices.info/jcms/int_68096/le-desequilibre-de-lapprentissage

9.4.2. Popular Science Radio Broadcast

France Culture (2012), interview of P-Y. Oudeyer, La robotique pour mieux comprendre l'homme, (Interview, 45 mn), Emission « Continent Sciences » de Stéphane Deligeorges. http://www.franceculture.fr/emission-continent-sciences-pierre-yves-oudeyer-2012-01-16
France Info (2012), interview of P-Y. Oudeyer, Robotique et Sciences Cognitives (Interview, 5 minutes). http://www.pyoudeyer.com/FranceInfo19Jan2012.mpg
RCF Aquitaine (october 2012) interview of Fabien Benureau. It allowed us to communicate on PhD work on autonomous intrinsic motivation algorithms done at our lab.

9.4.3. Popular Science Videos

Ly, O., Oudeyer, P-Y., Langlois, A. (2012) Le déséquilibre de l'apprentissage, Inria (selected in category "hors-compétition" at Festival du Film de Chercheur à Nancy, to be used as support in schools). http://interstices.info/jcms/int_68096/le-desequilibre-de-lapprentissage
Langlois, A., Oudeyer, P-Y. (2012) Alan Turing et la robotique développementale (interview of Pierre-Yves Oudeyer), Vidéothèque Inria. http://www.pyoudeyer.com/turing_oudeyer_inria_2012.mp4

9.4.4. Popular Science Talks

(4th august 2012) "Un robot peut-il apprendre comme un enfant?", Marathon des Sciences, Festival d'astronomie de Fleurance, Fleurance, France. http://www.festival-astronomie.com/
(30th january 2012), P-Y. Oudeyer: "Les modèles robotiques: un nouveau langage pour comprendre le vivant", Colloque "Mathématiques pour tous?", UNESCO, organisé par l'UNESCO en partenariat avec l'IHES et la Fondation Cartier pour l'Art Contemporain, UNESCO, Paris, France. http://www.science.gouv.fr/fr/agenda/bdd/res/4374/colloque-mathematiques-pour-tous-/
(14th december 2012), P-Y. Oudeyer, "Design et Auto-Design de Comportements et d'Interactions chez les Robots", Escales du Design, Bordeaux.

9.4.5. Museum exhibitions, science festivals and general public demonstrations

9.4.5.1. Ergo-Robots, exhibition "Mathematics, a Beautiful Elsewhere" at Fondation Cartier pour l'Art Contemporain

The FLOWERS team, in collaboration with University Bordeaux I/Labri, has participated as a central actor of the exhibition "Mathematics: A Beautiful Elsewhere" at Fondation Cartier pour l'Art Contemporain in Paris. This installation, called "Ergo-Robots/FLOWERS Fields" was made in collaboration with artist David Lynch and mathematician Mikhail Gromov (IHES, France), and shows computational models of curiosity-driven learning, human-robot interaction as well as self-organization of linguistic conventions. This exhibition, at the crossroads of science and art, allowed to disseminate our work towards the general public, explaining concepts related to learning mechanims in humans and robots to a large audience (80000 visitors). This was also an opportunity for experimenting and improving our technologies for life-long robot learning experimentation. For one of the first times in the world outside the laboratory, we demonstrated how it is possible to achieve experimentation with learning robots quasi-continously for 5 months. This opens novel stimulating scientific perspectives in the field of developmental robotics. This experimentation was presented through large audience radios, magazines and newspapers (France Inter, France Culture, RFI, Sciences et Avenir, Tangente, Financial Times, Daily Telegraph, Liberation, ...).
More information available at: http://flowers.inria.fr/ergo-robots.php and http://fondation.cartier.com/.

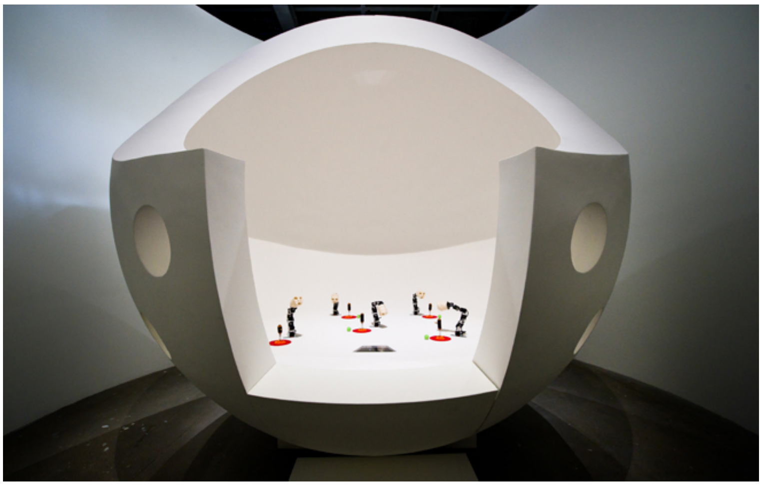


Figure 30. The Ergo-Robot experiment: robot learning experiment running continuously for 5 months at Fondation Cartier pour l'Art Contemporain, exhibition "Mathématiques: Un Depaysement Soudain".

9.4.5.2. Cap Sciences exhibition on "Brain and Cognition"

Cap Sciences is an organization in Bordeaux to promote and to communicate about science to the public. Cap Sciences is preparing an exhibit about the brain starting in February 2013. The Flowers team will contribute to this exposition by setting up a booth to explain the complexity of the processing required for intelligent artificial systems (e.g. robots) to transform observations from the environment to actions done in this environment, such processing being done continuously by all living beings, most notably by nervous systems and brains. To explain this idea, the Flowers team is working on a game for the visitors of the exhibit: a player has to drive forward a mobile robot, specifically an iRobot Roomba, while avoiding obstacles. The difficulty for the visitor in this game is that the player is not able to watch the robot in its environment: the player has to control it using only the sensory-information displayed on a computer screen (see figure 31). The player wins when the robot has traveled a given distance in a straight line and in limited time without bumping into an obstacle. This exhibit will start in February 2013 and last for a year. After that, it may move to different locations. More than 100,000 visitors are expected in Cap Sciences, half of them will come from elementary schools. The data generated by the robot and the visitors will be logged and will be available for research on life long learning with robots.

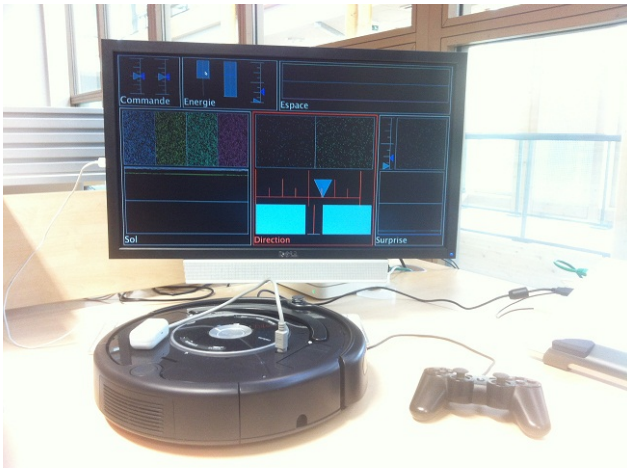


Figure 31. Picture of the mobile robot with the computer screen displaying the sensori-motor information the visitor needs to use to control the robot using a joystick.

9.4.5.3. Robots at International Exposition held in Yeosu, South Korea

. In collaboration with Rhoban project/Labri/CNRS/Univ. Bordeaux I, the Flowers team participated to a project where several robots were elaborated and installed at the May 12, 2012 - August 12 2012 International Exposition held in Yeosu, South Korea (600k visitors). Exhibited robots were three humanoids (one dancing, two playing on a spring) and five musicians (arms only) playing musical instruments (electric guitar, electric bass guitar, keytar, drums, DJ turntables). Robots were installed inside the french ward, in a specific room named botanic garden. Humanoids were closed to the audience to allow interaction between people and robots (see Figure 33) while musicians robots were higher on a dedicated wall to increase visibility of the show (see Figure 32).

9.4.5.4. Science Festivals:

- 13 october 2012 : Lab visit and robot demonstration for the 'Fete de la science' in the robotics lab at ENSTA ParisTech, Palaiseau.
- 9, 10 et 11 Février 2012: Aquitec exhibiton. Acquitec is an opportunity for high-school students to discover jobs and formations directly from the schools and institutes that exhibit there. We presented few robotic platforms. (Fabien Benureau and Jonathan Grizou)

9.4.6. Press

Web links to the following press items are available on http://flowers.inria.fr/press.php.

9.4.6.1. TV

Universcience TV, jan. 2012: "Art et maths".

France 2, Tele-matin, jan. 2012: "Les Maths: le 'probleme' des enfants".

Figure 32. Exhibited musician robots on the wall of the french ward

Figure 33. Exhibited humanoid robots on the wall of the french ward

France TV/Culturebox, jan. 2012: ”Mathematiques: Un Depaysement Soudain” A la Fondation Cartier pour l’Art Contemporain.

9.4.6.2. Radio

France Info, 19 jan. 2012: ”Robotique et sciences cognitives” (3mn).

France Culture, jan. 2012, Entretien sur le sujet ”La robotique pour mieux comprendre l’homme” (45 mn), Emission Continent Sciences de Stephane Deligeorges. A propos du contexte scientifique dans lequel s’inscrit le projet Ergo-Robots.

9.4.6.3. Magazines

Jan 2012: Sciences et Avenir Hors-Serie, Numero Special ”Qu’est-ce-que l’homme”: ”GX-29 n’est pas un objet comme les autres”.

9.4.6.4. Newspapers

Sud-Ouest, march, 2012: ”Les Etonnants robots de la Fondation Cartier”.

10. Bibliography

Major publications by the team in recent years

[1] A. ANGELI, D. FILLIAT, S. DONCIEUX, J. MEYER. *Fast and incremental method for loop-closure detection using bags of visual words*, in "Robotics, IEEE Transactions on", 2008, vol. 24, n^o 5, p. 1027–1037.

[2] A. BARANES, P.-Y. OUDEYER. *RIAC: Robust Intrinsically Motivated Exploration and Active Learning*, in "IEEE Trans. on Auto. Ment. Dev.", 2009, vol. 1, n° 3, p. 155-169, http://www.pyoudeyer.com/TAMDBaranesOudeyer09.pdf.

[3] A. BARANES, P.-Y. OUDEYER. *Active learning of inverse models with intrinsically motivated goal exploration in robots*, in "Robotics and Autonomous Systems", 2013, vol. 61, n° 1, p. 49 - 73 [*DOI :* 10.1016/J.ROBOT.2012.05.008], http://www.pyoudeyer.com/RAS-SAGG-RIAC-2012.pdf.

[4] J. BUCHLI, F. STULP, E. THEODOROU, S. SCHAAL. *Learning Variable Impedance Control*, in "International Journal of Robotics Research", 2011, vol. 30, n° 7, p. 820-833, http://ijr.sagepub.com/content/early/2011/03/31/0278364911402527.

[5] T. DEGRIS, O. SIGAUD, P. WUILLEMIN. *Learning the Structure of Factored Markov Decision Processes in Reinforcement Learning Problems*, in "Proceedings of the 23rd International Conference on Machine learning (ICML)", 2006, p. 257–264.

[6] T. DEGRIS, M. WHITE, R. SUTTON. *Off-Policy Actor-Critic*, in "International Conference on Machine Learning", 2012, http://hal.inria.fr/hal-00764021.

[7] D. FILLIAT. *A visual bag of words method for interactive qualitative localization and mapping*, in "Robotics and Automation, 2007 IEEE International Conference on", IEEE, 2007, p. 3921–3926.

[8] A. GEPPERTH. *Efficient online bootstrapping of sensory representations*, in "Neural Networks", December 2012 [*DOI :* 10.1016/J.NEUNET.2012.11.002], http://hal.inria.fr/hal-00763660.

[9] A. GEPPERTH, S. REBHAN, S. HASLER, J. FRITSCH. *Biased competition in visual processing hierarchies: a learning approach using multiple cues*, in "Cognitive Computation", March 2011, vol. 3, n^o 1, http://hal.archives-ouvertes.fr/hal-00647809/en/.

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

[11] M. LOPES, F. MELO, L. MONTESANO. *Active learning for reward estimation in inverse reinforcement learning*, in "Machine Learning and Knowledge Discovery in Databases", 2009, p. 31–46.

[12] L. MONTESANO, M. LOPES, A. BERNARDINO, J. SANTOS-VICTOR. *Learning Object Affordances: From Sensory–Motor Coordination to Imitation*, in "Robotics, IEEE Transactions on", 2008, vol. 24, n^o 1, p. 15–26.

[13] S. M. NGUYEN, A. BARANES, P.-Y. OUDEYER. *Bootstrapping Intrinsically Motivated Learning with Human Demonstrations*, in "proceedings of the IEEE International Conference on Development and Learning", Frankfurt, Allemagne, 2011, ERC Grant EXPLORERS 240007, http://hal.archives-ouvertes.fr/hal-00645986.

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

[15] P.-Y. OUDEYER. *Self-Organization in the Evolution of Speech*, Studies in the Evolution of Language, Oxford University Press, 2006.

[16] P.-Y. OUDEYER. *On the impact of robotics in behavioral and cognitive sciences: from insect navigation to human cognitive development*, in "IEEE Transactions on Autonomous Mental Development", 2010, vol. 2, n^o 1, p. 2–16, http://hal.inria.fr/inria-00541783/en/.

[17] P. ROUANET, P.-Y. OUDEYER, F. DANIEAU, D. FILLIAT. *The Impact of Human-Robot Interfaces on the Learning of Visual Objects*, in "IEEE Transactions on Robotics", January 2013, http://hal.inria.fr/hal-00758241.

[18] F. STULP, B. BUCHLI, A. ELLMER, M. MISTRY, E. THEODOROU, S. SCHAAL. *Model-free Reinforcement Learning of Impedance Control in Stochastic Force Fields*, in "IEEE Transactions on Autonomous Mental Development", 2012.

[19] F. STULP, A. FEDRIZZI, L. MÖSENLECHNER, M. BEETZ. *Learning and Reasoning with Action-Related Places for Robust Mobile Manipulation*, in "Journal of Artificial Intelligence Research (JAIR)", 2012, vol. 43, p. 1–42.

[20] F. STULP, E. THEODOROU, S. SCHAAL. *Reinforcement Learning with Sequences of Motion Primitives for Robust Manipulation*, in "IEEE Transactions on Robotics", 2012, vol. 28, n^o 6, p. 1360-1370.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[21] P. ROUANET. *Apprendre à un robot à reconnaître des objets visuels nouveaux et à les associer à des mots nouveaux : le rôle de l'interface*, Université Sciences et Technologies - Bordeaux I, April 2012, http://hal.inria.fr/tel-00758249.

Articles in International Peer-Reviewed Journals

[22] A. BARANES, P.-Y. OUDEYER. *Active learning of inverse models with intrinsically motivated goal exploration in robots*, in "Robotics and Autonomous Systems", 2013, vol. 61, n° 1, p. 49 - 73 [*DOI :* 10.1016/J.ROBOT.2012.05.008], http://www.pyoudeyer.com/RAS-SAGG-RIAC-2012.pdf.

[23] A. GEPPERTH, B. DITTES, M. GARCIA ORTIZ. *The contribution of context information: a case study of object recognition in an intelligent car*, in "Neurocomputing", February 2012, vol. 94 [*DOI :* 10.1016/J.NEUCOM.2012.03.008], http://hal.inria.fr/hal-00763650.

[24] A. GEPPERTH. *Efficient online bootstrapping of sensory representations*, in "Neural Networks", December 2012 [*DOI :* 10.1016/J.NEUNET.2012.11.002], http://hal.inria.fr/hal-00763660.

[25] C. MOULIN-FRIER, R. LAURENT, P. BESSIÈRE, J.-L. SCHWARTZ, J. DIARD. *Adverse conditions improve distinguishability of auditory, motor and percep-tuo-motor theories of speech perception: an exploratory Bayesian modeling study*, in "Language and Cognitive Processes", 2012, vol. 27, n° 7-8 Special Issue: Speech Recognition in Adverse Conditions, p. 1240-1263 [*DOI :* 10.1080/01690965.2011.645313], http://hal.archives-ouvertes.fr/hal-00642311.

[26] P. PASTOR, M. KALAKRISHNAN, F. MEIER, F. STULP, J. BUCHLI, E. THEODOROU, S. SCHAAL. *From Dynamic Movement Primitives to Associative Skill Memories*, in "Robotics and Autonomous Systems", 2012.

[27] P. PILARSKI, M. DAWSON, T. DEGRIS, J. CAREY, K. CHAN, J. HEBERT, R. SUTTON. *Prediction and Anticipation for Adaptive Artificial Limbs*, in "IEEE Robotics and Automation Magazine. Special Issue on Assistive Robotics", 2013.

[28] P. ROUANET, P.-Y. OUDEYER, F. DANIEAU, D. FILLIAT. *The Impact of Human-Robot Interfaces on the Learning of Visual Objects*, in "IEEE Transactions on Robotics", January 2013, http://hal.inria.fr/hal-00758241.

[29] F. STULP, B. BUCHLI, A. ELLMER, M. MISTRY, E. THEODOROU, S. SCHAAL. *Model-free Reinforcement Learning of Impedance Control in Stochastic Force Fields*, in "IEEE Transactions on Autonomous Mental Development", 2012.

[30] F. STULP, A. FEDRIZZI, L. MÖSENLECHNER, M. BEETZ. *Learning and Reasoning with Action-Related Places for Robust Mobile Manipulation*, in "Journal of Artificial Intelligence Research (JAIR)", 2012, vol. 43, p. 1–42.

[31] F. STULP, E. THEODOROU, S. SCHAAL. *Reinforcement Learning with Sequences of Motion Primitives for Robust Manipulation*, in "IEEE Transactions on Robotics", 2012, vol. 28, n° 6, p. 1360-1370.

International Conferences with Proceedings

[32] M. CAKMAK, M. LOPES. *Algorithmic and Human Teaching of Sequential Decision Tasks*, in "AAAI Conference on Artificial Intelligence (AAAI-12)", Toronto, Canada, July 2012, http://hal.inria.fr/hal-00755253.

[33] A. CHAPOULIE, P. RIVES, D. FILLIAT. *Topological segmentation of indoors/outdoors sequences of spherical views*, in "IEEE Conference on Intelligent Robots and Systems, IROS'12", Vilamoura, Portugal, October 2012, p. 4288-4295, http://hal.inria.fr/hal-00752909.

[34] A. CSAPO, E. GILMARTIN, J. GRIZOU, J. HAN, R. MEENA, D. ANASTASIOU, K. JOKINEN, G. WILCOCK. *Multimodal Conversational Interaction with a Humanoid Robot*, in "CogInfoCom 2012 (3rd IEEE International Conference on Cognitive Infocommunications)", Kosice, Slovakia, December 2012, http://hal.inria.fr/hal-00759810.

[35] T. DEGRIS, J. MODAYIL. *Scaling-up Knowledge for a Cognizant Robot*, in "AAAI Spring Symposium on Designing Intelligent Robots: Reintegrating AI.", Stanford, United States, March 2012, http://hal.inria.fr/hal-00764289.

[36] N. DEGRIS, P. PILARSKI, R. SUTTON. *Apprentissage par Renforcement sans Modèle et avec Action Continue*, in "Journées Francophones sur la planification, la décision et l'apprentissage pour le contrôle des systèmes - JFPDA 2012", Villers-lès-Nancy, France, O. BUFFET (editor), 2012, 11 p, http://hal.inria.fr/hal-00736314.

[37] T. DEGRIS, P. PILARSKI, R. SUTTON. *Model-Free Reinforcement Learning with Continuous Action in Practice*, in "American Control Conference", Montreal, Canada, June 2012, http://hal.inria.fr/hal-00764281.

[38] T. DEGRIS, M. WHITE, R. SUTTON. *Off-Policy Actor-Critic*, in "International Conference on Machine Learning", Edinburgh, United Kingdom, June 2012, http://hal.inria.fr/hal-00764021.

[39] D. FILLIAT, E. BATTESTI, S. BAZEILLE, G. DUCEUX, A. GEPPERTH, L. HARRATH, I. JEBARI, R. PEREIRA, A. TAPUS, C. MEYER, S.-H. IENG, R. BENOSMAN, E. CIZERON, J.-C. MAMANNA, B. POTHIER. *RGBD object recognition and visual texture classification for indoor semantic mapping*, in "Technologies for Practical Robot Applications (TePRA), 2012 IEEE International Conference on", United States, 2012, p. 127 - 132, http://hal.inria.fr/hal-00755295.

[40] A. GEPPERTH, L.-C. CARON. *Simultaneous concept formation driven by predictability*, in "International conference on development and learning", San Diego, États-Unis, October 2012.

[41] A. GEPPERTH. *Co-training of context models for real-time object detection*, in "IEEE Symposium on Intelligent Vehicles", Madrid, Espagne, June 2012, http://hal.inria.fr/hal-00763676.

[42] T. HESTER, M. LOPES, P. STONE. *Learning Exploration Strategies in Model-Based Reinforcement Learning*, in "AAMAS", 2013.

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

[44] J. KULICK, M. TOUSSAINT, T. LANG, M. LOPES. *Active Learning for Teaching a Robot Grounded Relational Symbols*, in "ICRA", 2013, submitted.

[45] C. LOPERA, H. TOMÉ, A. R. TSOUROUKDISSIAN, F. STULP. *Comparing Motion Generation and Motion Recall for Everyday Robotic Tasks*, in "12th IEEE-RAS International Conference on Humanoid Robots", 2012.

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

[47] M. LOPES, P.-Y. OUDEYER. *The Strategic Student Approach for Life-Long Exploration and Learning*, in "IEEE Conference on Development and Learning / EpiRob 2012", San Diego, United States, November 2012, http://hal.inria.fr/hal-00755216.

[48] N. LYUBOVA, D. FILLIAT. *Developmental Approach for Interactive Object Discovery*, in "Neural Networks (IJCNN), The 2012 International Joint Conference on", Australia, 2012, p. 1-7, http://hal.inria.fr/hal-00755298.

[49] N. LYUBOVA, D. FILLIAT. *Developmental Learning for Object Perception*, in "Proceedings of the CogSys2012 Workshop on Deep Hierarchies in Vision", Autriche, 2012, http://hal.archives-ouvertes.fr/hal-00755300.

[50] A. MAHMOOD, R. SUTTON, T. DEGRIS, P. PILARSKI. *Tuning-Free Step-Size Adaptation*, in "Proceedings of the 37th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)", 2012, p. 2121–2124.

[51] O. MANGIN, P.-Y. OUDEYER. *Learning the Combinatorial Structure of Demonstrated Behaviors with Inverse Feedback Control*, in "Human Behavior Understanding", Springer, October 2012, vol. 7559, Lecture notes in computer science, http://hal.inria.fr/hal-00764448.

[52] O. MANGIN, P.-Y. OUDEYER. *Learning to recognize parallel combinations of human motion primitives with linguistic descriptions using non-negative matrix factorization*, in "Proceedings of 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems", IEEE, October 2012, http://hal.inria.fr/hal-00764353.

[53] *Best Paper*
C. MOULIN-FRIER, P.-Y. OUDEYER. *Curiosity-driven phonetic learning*, in "ICDL-Epirob - International Conference on Development and Learning, Epirob", San Diego, États-Unis, November 2012, http://hal.inria.fr/hal-00762795.

[54] S. M. NGUYEN, P.-Y. OUDEYER. *Interactive Learning Gives the Tempo to an Intrinsically Motivated Robot Learner*, in "IEEE-RAS International Conference on Humanoid Robots", Osaka, Japan, December 2012, http://hal.inria.fr/hal-00762753.

[55] S. M. NGUYEN, P.-Y. OUDEYER. *Properties for Efficient Demonstrations to a Socially Guided Intrinsically Motivated Learner*, in "21st IEEE International Symposium on Robot and Human Interactive Communication", Paris, France, September 2012, http://hal.inria.fr/hal-00762758.

[56] S. M. NGUYEN, P.-Y. OUDEYER. *Whom Will an Intrinsically Motivated Robot Learner Choose to Imitate from?*, in "Post-Graduate Conference on Robotics and Development of Cognition: RobotDoC-PhD 2012", Lausanne, Switzerland, September 2012, http://hal.inria.fr/hal-00762762.

[57] P. PILARSKI, M. DAWSON, T. DEGRIS, J. CAREY, R. SUTTON. *Dynamic switching and real-time machine learning for improved human control of assistive biomedical robots*, in "Proceedings of the International Conference on Biomedical Robotics and Biomechatronics (BioRob), 4th IEEE RAS EMBS", 2012, p. 296 -302 [*DOI :* 10.1109/BIOROB.2012.6290309].

[58] P. PILARSKI, T. DEGRIS, M. DAWSON, J. CAREY, K. CHAN, J. HEBERT, R. SUTTON. *Towards Prediction-Based Prosthetic Control*, in "Proceedings of the 17th International Functional Electrical Stimulation Society Conference (IFESS)", 2012.

[59] A. A. SALAH, J. R. DEL SOLAR, Ç. MERIÇLI, P.-Y. OUDEYER. *Human Behavior Understanding for Robotics.*, in "Human Behavior Understanding", A. A. SALAH, J. R. DEL SOLAR, Ç. MERIÇLI, P.-Y. OUDEYER (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7559, p. 1-16.

[60] F. STULP, P.-Y. OUDEYER. *Emergent Proximo-Distal Maturation through Adaptive Exploration*, in "International Conference on Development and Learning (ICDL)", 2012.

[61] F. STULP, O. SIGAUD. *Adaptation de la matrice de covariance pour l'apprentissage par renforcement direct*, in "Journées Francophones sur la planification, la décision et l'apprentissage pour le contrôle des systèmes - JFPDA 2012", Villers-lès-Nancy, France, O. BUFFET (editor), 2012, 12 p, http://hal.inria.fr/hal-00736310.

[62] F. STULP, O. SIGAUD. *Path Integral Policy Improvement with Covariance Matrix Adaptation*, in "Proceedings of the 29th International Conference on Machine Learning (ICML)", 2012.

[63] F. STULP. *Adaptive Exploration for Continual Reinforcement Learning*, in "International Conference on Intelligent Robots and Systems (IROS)", 2012.

[64] B. WREDE, K. J. ROHLFING, J. J. STEIL, S. WREDE, P.-Y. OUDEYER, J. TANI. *Towards robots with teleological action and language understanding*, in "IEEE Humanoids, Workshop on Developmental Robotics", Osaka, IEEE, 2012, http://www.pyoudeyer.com/Wredeetal12.pdf.

National Conferences with Proceeding

[65] T. DEGRIS, M. PILARSKI, R. SUTTON. *Apprentissage par Renforcement sans Modèle et avec Action Continue*, in "7èmes Journées Francophones Planification, Décision, et Apprentissage pour la conduite de systèmes", Nancy, France, May 2012, http://hal.inria.fr/hal-00764325.

Scientific Books (or Scientific Book chapters)

[66] T. CEDERBORG, P.-Y. OUDEYER. *Learning words by imitating*, in "Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence", L. GOGATE, G. HOLLICH (editors), IGI Global, January 2012.

Books or Proceedings Editing

[67] A. A. SALAH, J. R. DEL SOLAR, Ç. MERIÇLI, P.-Y. OUDEYER. , A. A. SALAH, J. R. DEL SOLAR, Ç. MERIÇLI, P.-Y. OUDEYER (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7559.

Other Publications

[68] R. LAURENT, C. MOULIN-FRIER, P. BESSIÈRE, J.-L. SCHWARTZ, J. DIARD. *Integrate, yes, but what and how? A computational approach of perceptuo-motor fusion in speech perception.*, 2012, Accepted commentary in Behavioral and Brain Sciences, in Press., http://hal.inria.fr/hal-00765973.

[69] F. STULP, O. SIGAUD. *Policy Improvement Methods: Between Black-Box Optimization and Episodic Reinforcement Learning*, 2012, 34 pages, http://hal.inria.fr/hal-00738463.

References in notes

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

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

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

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

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

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

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

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

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

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

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

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

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

[83] T. DEGRIS, M. WHITE, R. SUTTON. *Off-Policy Actor-Critic*, in "arXiv preprint arXiv:1205.4839", 2012.

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

[85] C. FLEISCHER, A. WEGE, K. KONDAK, G. HOMMEL. *Application of EMG signals for controlling exoskeleton robots*, in "Biomedizinische Technik", 2006, vol. 51, n^o 5/6, p. 314–319.

[86] D. GOUAILLIER, V. HUGEL, P. BLAZEVIC, C. KILNER, J. MONCEAUX, P. LAFOURCADE, B. MARNIER, J. SERRE, B. MAISONNIER. *The nao humanoid: a combination of performance and affordability*, in "CoRR, vol. abs/0807.3223", 2008.

[87] I. HA, Y. TAMURA, H. ASAMA, J. HAN, D. HONG. *Development of open humanoid platform DARwIn-OP*, in "SICE Annual Conference (SICE), 2011 Proceedings of", IEEE, 2011, p. 2178–2181.

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

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

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

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

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

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

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

[95] M. LAPEYRE, O. LY, P. OUDEYER. *Maturational constraints for motor learning in high-dimensions: the case of biped walking*, in "Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on", IEEE, 2011, p. 707–714.

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

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

[98] O. LY, M. LAPEYRE, P. OUDEYER. *Bio-inspired vertebral column, compliance and semi-passive dynamics in a lightweight humanoid robot*, in "Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on", IEEE, 2011, p. 1465–1472.

[99] O. MANGIN, P.-Y. OUDEYER. *Unsupervised learning of simultaneous motor primitives through imitation*, in "IEEE ICDL-EPIROB 2011", Frankfurt, Germany, August 2011, http://hal.inria.fr/hal-00652346/en.

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

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

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

[103] L. MONTESANO, M. LOPES. *Active learning of visual descriptors for grasping using non-parametric smoothed beta distributions*, in "Robotics and Autonomous Systems", 2011, p. 26-AUG-2011 [*DOI :* 10.1016/J.ROBOT.2011.07.013], http://hal.inria.fr/hal-00637575/en.

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

[105] S. M. NGUYEN, A. BARANES, P.-Y. OUDEYER. *Bootstrapping Intrinsically Motivated Learning with Human Demonstrations*, in "proceedings of the IEEE International Conference on Development and Learning", Frankfurt, Germany, 2011, ERC Grant EXPLORERS 240007, http://hal.archives-ouvertes.fr/hal-00645986.

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

[107] R. NOWAK. *The Geometry of Generalized Binary Search*, in "Information Theory, Transactions on", 2011, vol. 57, n^{o} 12, p. 7893–7906.

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

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

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

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

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

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

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

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

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

[117] M. SCHWARZ, M. SCHREIBER, S. SCHUELLER, M. MISSURA, S. BEHNKE. *NimbRo-OP Humanoid TeenSize Open Platform*.

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

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

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

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

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