# Maturational constraints for motor learning in high-dimensions: the case of biped walking

Matthieu Lapeyre, Olivier Ly, Pierre-Yves Oudeyer

# Maturational constraints for motor learning in high-dimensions: the case of biped walking

Matthieu Lapeyre
Flowers Research Team
INRIA Bordeaux Sud-Ouest
Talence, France
matthieu.lapeyre@inria.fr

Olivier Ly
Flowers Research Team/ Labri
INRIA Bordeaux Sud-Ouest/ University Bordeaux I
Talence, France
ly@labri.fr

Pierre-Yves Oudeyer
Flowers Research Team
INRIA Bordeaux Sud-Ouest
Talence, France
pierre-yves.oudeyer@inria.fr

***Abstract*—This paper outlines a new developmental approach to motor learning in very high-dimensions, applied to learning biped locomotion in humanoid robots. This approach relies on the formal modeling and coupling of several advanced mechanisms inspired from human development for actively controlling the growth of complexity and harnessing the curse of dimensionality: 1) Maturational constraints for the progressive release of new degrees of freedoms and progressive increase their explorable ranges; 2) Motor synergies; 3) Morphological computation; 4) Social Guidance. An experimental setup involving a simulated version of the Acroban Humanoid robot is presented.**

**Keywords: Robot Learning, Neuro-robotics**

## I. Introduction

For robots and especially humanoid robots, one of the most important challenges today and in the future is to be robust and reactive to unpredicted events. Indeed, robots will be more and more confronted to a public environment where the world is highly unpredictable and changing. Also, it seems impossible that engineers can provide an effective behavior for every situation and every environment that the robot will encounter during its life. So, in order to interact in physical and social environments which are initially unknown and changing, the robots should find a correct behaviour by itself or through natural social interaction. One way is to implement mechanisms that allow robots to learn new skills and adapt them along their whole life.

In this paper, our study investigates the challenges raised by motor learning in high dimensions. Indeed, including its own body and the open-ended surrounding physical and social environment, the continuous sensorimotor space of a typical robot is extremely large and high dimensional, raising a major problem for the learning of new skills. For such high dimensional and unbounded spaces, a random sampling without special constraints, even with a relatively fast simulation, can not be practiced and learnt within a lifetime.

This is especially the case for humanoids robots requiring coordinated movements of the whole body, which mostly include at least 30 actuators. In addition, each actuator is typically controlled and characterized by several parameters including position, speed, acceleration, this increase of dimensionality provides novel important conceptual and technical challenges for motor learning. In particular, concerning the acquisition of biped walking, i.e. walking without static equilibrium, which is often considered as an extremely ambitious challenge [1], [2]. Trying to pre-program biped dynamic walking through classical engineering methods is equally difficult requiring highly complex and precise model of the robot's mechanic dynamic [3], [4].

Even using advanced gradient-based optimization techniques learning without constraints would fail given the very high ruggedness of the fitness landscape corresponding to the dynamics of a humanoid with many degree of freedom and foot contact. Also, the dynamics of the physics of real world whole-body humanoid robot is so complicated that even the most advanced approaches for designing analytical controllers by computation has not yet produced controllers able to allow many state-of-the art robots (e.g. Hondas ASIMO or Kawada Industries HRP-4 [3], [4]) to be robust to unpredicted perturbations of the real world (e.g. obstacles on the ground or a human giving a tap in the back). Thus, walking has all the apparent features of a skill that is very difficult both to learn and/or engineer. Humans have a much more complex body than the most advanced robots, involving hundreds of joints and thousand of sensors. However, the can acquire new skills in a faster, safer and more robust way than robots and in especially concerning the dynamic walking. Millions years of evolution have brought some exploration strategies, mechanisms and constraints in the human learning system in order to speed up and reduce the complexity of learning new skills even in a very high dimensional sensorimotor space. In our work, we call those mechanisms *"developmental constraints"* in reference of the development of human infant. The main hypothesis of our work is that developmental constraints should be introduced in order to reduce and constrain the growth of the size and complexity of practically explorable spaces. Those mechanisms should essentially allow the organism to automatically introduce self-bounding in the unbounded world (including their own body), and then progressively releasing constraints and bounding to increase the volume and the dimensionality of explorable sensorimotor spaces, i.e. the diversity of explorable knowledge and skills. Most of these developmental constraints that we are investigating are strongly inspired by constraints on human infant development from which we take the fundamental insight. The complex acquisition of novel skills in the real world necessitates sophisticated innate capabilities/constraints that may unfold with time in interaction with the environment during the course of epigenesis [5].

In the following, we will describe some of them and explain how they may facilitate, sometimes considerably, the exploration and acquisition of complex skills in real-world sensorimotor spaces, more precisely:

- Parameterized sensori and motor primitives, also referred as muscle synergies controlled by neural dynamical systems;
- Embodiment and morphological computation;
- Maturational constraints;
- Social guidance.

In this paper, we propose to study some developmental constraints found in human development in order to formalized a developmental learning algorithm. This algorithm uses maturational constraints in order to reduce the motor space learnable space and social assistance to guide the robot. Then, we will experiment this algorithm with a setup involving our robot Acroban[1]. This algorithm will be applied to the learning of bipedal walking.

## II. General review of Developmental Constraints

### A. Sensorimotor primitives

First, human babies are born with neurally embedded dynamical systems which on the sensory side allow them to be able to detect and track a number of higher-level structures right from the start, and on the motor side allow them to tune motor and muscle synergies which already generate parameterized coordinated movements [6], [7], [8]. Examples of motor primitives include central pattern generators such as for leg oscillations [9] or synergies for reaching with the hand (e.g. [10]).

Those primitives are typically parameterized, and thus can typically be seen as parameterized dynamical systems which semantics (affordances in particular), parameter values to be set and combination for achieving given tasks have to be learnt. For example, central pattern generators are typically neurally implemented as complex dynamical system generating oscillatory movements, which can be tuned by controlling a number of high-level parameters. Yet, these sensorimotor primitives can considerably decrease the dimensionality, and thus the size of the explorable sensorimotor spaces and transform complex low-level action planning problems in simpler higher-level dynamical system tuning problems. In the case of human locomotion, the walking behavior depends on an extension-flexion rhythmic generator of the limb controlled by pre-wired networks responsible for the gait cycle. The existence of these rhythmic oscillators and multi-joint integrated units that control the activity of legs has now been demonstrated at the spinal level [11]. Adjusting the rate of rotation is obtained simply by varying the intensity of this command. The oscillator coordinates the various joint modules composing each legs and it expresses its rhythmic commands independently of any sensory feedback, i.e. open loop [12].

[1]Humanoid robot with large number of articulation. It will be described later in this paper.

### B. Embodiment and morphological computation

In many robots, the morphology or the design is not correlated with the controller. However, animals show a high dependency between morphology and control, as the salamander for exemple[13]. Indeed, the efficiency of those primitives is tightly related to the morphological properties of the body in which they are used. Also, the inputs and structure of those primitives only make sense within a given body structure. The outputs of those primitives do not entirely determine the movements/behaviour of the robot body. The physics of real-world robots is such that gravity and its interaction with the inertia of the robot, in combination with the compliance of materials and actuators, also importantly impacts the resulting movements/behaviour. Thus, the impact of morphology on control and behaviour is paramount. An adequately designed morphology can allow to significantly reduce the complexity of its traditional control code/system for a given set of tasks, and can even be conceptualized as replacing traditional digital control computations by physical or morphological computation [14], [15], [16]. The body itself, as a physical dynamical system subject to the laws of physics, should actually be considered as any other complex dynamical system, which can potentially generate spontaneously organized structures through self-organization [17].

### C. Maturational constraints

The challenge of motor learning in high-dimensions is typically associated with robots using a large number of degree of freedom. This is especially the case of humanoid robot involving many actuators, but its above all the case of human whose anatomy is highly complex and nonlinear, composed of more than 300 articulations and 600 muscles potentially redundant. As argued at the beginning of this article, mechanisms for self-bounding the explorable space are necessary, but they should be as little ad hoc as possible. To reach this objective, one may take inspiration from maturational mechanisms in biological organisms. A few telling examples of constraints in the sensory, motor and neural systems of vertebrate species such as rats,cats and humans are the immaturity of the accommodative system [18], the low acuity of vision and absence of binocularity [19], the low leg muscle: leg fat ratio, and the poor postural control of head, trunk, arms and legs [20];[21].

The progressive biological maturation of infants brain, motor and sensor capabilities, introduces numerous important constraints on the learning process [Schlesinger, 2008]. Indeed, at birth, all the sensorimotor apparatus is neither precise enough nor fast enough, to allow infants to perform complex tasks. The degree of freedom problem was suggested by Bernstein (1967). He proposed that three steps existed in children and gradually release new degrees of freedom. First, when infants learn new skill (reaching, touching, walking ), articulations which are farther from the trunk, such as ankle and wrist are reduced to a minimum i.e. freezed. Then, as the infants progress in their learning, restrictions at the periphery are gradually lifted (freeing), until all degrees of freedom are incorporated. Eventually, reactive phenomena (such as gravity and passive dynamics) are exploited, and the most efficient movements are selected [22]. In a study on the pendulation of small-sized humanoid robot, Lungarella and Berthouze provided experimental evidence "that starting with fewer degrees of freedom enables a more efficient exploration of the sensorimotor space during the acquisition of a task"[23],[24]. Other examples, including adults learning a ski-simulator task [25], learning of a hand writing signature with the non-dominant limb [26], shown that subjects froze many joints of the whole body before introduce new active joints in a manner consistent with the Bernsteins theory of freeing degree of freedom. All these studies show that maturational constraints play an important role in learning, by partially determining a developmental pathway.

### D. Social/environnement guidance

In the weeks prior to independent walking, infants exhibit several transient upright skills that mitigate the requirements of single limb support. They hold furniture and pull up to a vertical position, stand while holding onto furniture, take forward steps while holding a caregivers hands, and cruise sideways in an upright position while holding onto furniture [27]. Each of these skills involves manual support of upright posture. The furniture or caregiver compensates for the missing levels of leg strength and balance control.

## III. Acroban

### A. A bio-inspired Morphology

The Acroban platform (see figure 1), more largely presented in [28] is a small (about 70cm) and lightweight (about 5kg) bio-inspired and compliant humanoid robot with many degrees of freedom (30 dofs) and a multiarticulated spine. The structure only includes revolute joints,which are all actuated by servomotors in a modular way:

- Each ankle has 3 joints enforcing a spherical link,
- Each knee has 1 joint enforcing a revolute link,
- Each hip has 3 joints enforcing a spherical link,
- The vertebral column has 5 joints,
- Each shoulder has 1 joint enforcing a 2-revolute joints link,
- Each elbow has 2 joints enforcing a 2-revolute joints link.

We essentially focused on designing a mechanically rich and open structure in the area of the vertebral column and the pelvis, providing it with 11 degrees of freedom on those areas.

*1) Vertebral column:* The vertebral column can be viewed as a system linking the pelvis and the shoulder. It enforces two revolutes joints links at its two extremities, each of them providing rotations in the sagittal and the coronal planes and one in the transverse plane. During motions, and in particular motions related to locomotion, this allows getting independency of the higher part and the lower part of the body. This allows for instance to reduce the dynamic of the higher part of the body during the gait. We claim that this contributes significantly to the stabilisation of the robot.

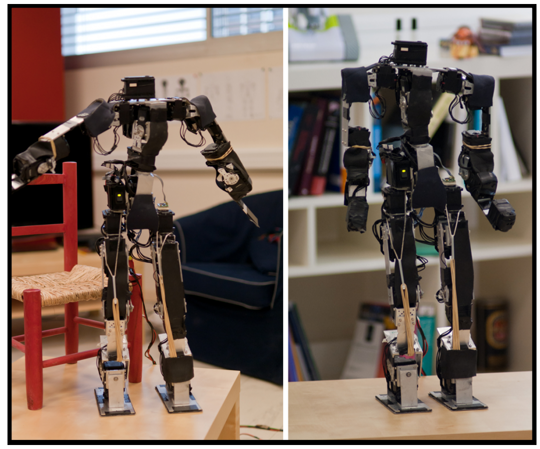

Fig. 1. The Acroban platform: a bio-inspired and compliant humanoid robot (more informations and videos on our website: http://flowers.inria.fr/acroban.php)

*2) Pelvis:* The pelvis, seen as an independent sub-body, may have several kind of mobility. It produces precise movements of the center of gravity of the robot. Firstly, it can move by a rotation in the sagittal plane. We will use it extensively to keep balance. Secondly, it can move in the transverse plane, this is used for the gait for the weight transfer between the legs, instead of making the legs support all the efforts of displacement of the body.

*3) Bio-inspiration of the mechanical structure:* The Human gait is an undeniable reference for the study of locomotion. Even if it is far from being clear that a direct transposition of human gait to robots is really effective, the mechanical and control processes generated by thousands years of evolution to solve bipedal locomotion problem are an important source of information and inspiration for humanoid robotics [29]. Many researchers were interested in biomechanics of human walking (see e.g [30]). These studies describe accurately the kinematics and dynamics of legs during walking (see [31]). Among the large litterature concerning human biped walking, only a few projects studied the role of the trunk during walking. Yet the trunk represents 60% of the total weight for humans, which raises the center of gravity (see e.g. [32], [33]). The trunk has a large complex network of muscles used to accomplish a lot of movements while keeping the balance. Its movements are regulated by a complex combinaison of anticipatory and reactive actions. The movements of the spine can facilitate the transfer of weight from one leg to the other one, improve the balance but also participate to the dynamic of the walking. It seems therefore interesting to enable a humanoid robot trying to walk in a robust way, to have an articulated trunk. But the human trunk is difficult to replicate on a small robot using servomotors. So we must simplify and find the most essential degrees of freedom of the spine. Ceccato [34] studied the role of the trunk and highlighted the main displacements of the spine during walking. And the apparent high dimensionality of the trunk appears to be factorizable down to a few essential components/dimensions. First, experiments showed small oscillations in the pelvis and the thoracic in the sagittal plane, highlighting that only two joints, one for the pelvis and one for the thoracic should be sufficient to represent the motion of the spine in the sagittal plane. In the coronal plane, the pelvis and shoulders oscillate in phase opposition while the middle remains straight throughout the cycle. This implies that essential movements of the trunk in the coronal plane could be approximated using two joints, one for the pelvis and an other one for the shoulder. Finally, in the horizontal plane, there are opposite rotations between the upper trunk and the lower trunk, enforced by a twist of the spine. So, only one revolute joint in the middle of the spine should be sufficient. Accordingly, Acroban has five joints for the trunk, as shown on figure 1 (see also Video 1): Two in the sagittal plane and two in the coronal plane, placed in the pelvis and shoulder/thoracic and one in the horizontal plane placed in the middle of his trunk. With this design, we have the strictly minimum necessary joints to replicate essential degrees of freedom of the human trunk.

### *B. Physical 3D-model of Acroban*

For the moment, experiments are done in a physical simulator (www.v-rep.eu). Indeed, we want to largely explore the impacts of developmental constraints on learning the bipedal walking, an important number of trials is required. As a first step, using simulators allow us to try faster and in parallel different solutions without the technical issue such as installing hardware or failure. The figure 2 shows a view of the Acroban V-rep model. To reduce computation time, the aspect of the robot is simplified but the physical properties (dimensions, mass, inertia) are kept.

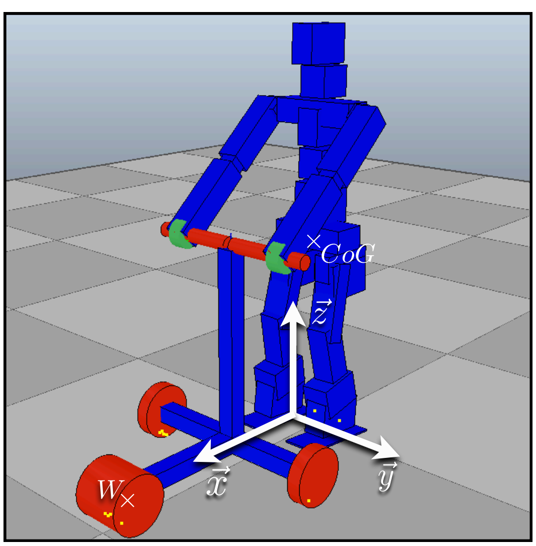


Fig. 2. Modelisation of Acroban in V-rep with axis orientation

## IV. Formalization

In this section, we propose to model each developmental constraints we developed in the previous section. These models will allow us to experiment on a virtual robot defined as:

Let us consider a robotic system, whose configurations/states are described in both an actuator space $\mathcal{J}$ and an operational space $\mathcal{W}$. For a given configurations $(j_1, w_1) \in \mathcal{J} \times \mathcal{W}$, $a$ sequence of actions $a = (a_1, a_2, a_3 a_n)$ allows a transition toward the new states $(j_1, w_2) \in \mathcal{J} \times \mathcal{W}$ such that $(j_1, w_1, a) \; => \; (w_2)$. In the case of a humanoid robot, $\mathcal{J}$ may represent its actuator/joint space, $\mathcal{W}$ the operation space corresponding to the cartesian position of the body Center of Gravity (CoG) in the world reference and $a$ may be the time position commands of each joints.

Our maturational constrained learning model considers distance reached $w_{CoG} \in \mathcal{W}$ at the end of simulation for a given set of action $a \in \mathcal{A}$ in a given actuator space $j \in \mathcal{J}$. The simulation ended when the robot falls (i.e. $w_{CoG}$ altitude is under a threshold) or when the simulation time reaches the maximum simulation time allowed $t_{end}$. The learning objective is defined as find $(j_{sol}, a_{sol}) \in \mathcal{J} \times \mathcal{A}$ such as maximizing the $x$ position of CoG (i.e go forward) at the end of simulation time.

It is necessary to find which maturational constraints are required to increase progressively the exploration space of motor primitives (defining a developmental pathway) to allow the robot to go as far as possible within the time $t_{end}$

### *A. Developmental Constraints*

*1) Motor Primitives:* In our work, the humanoid robot has 17 actuated degrees of freedom (see figure 3). In this way, the whole body of Acroban can be controlled excepted arms and its head. Arms are passive and hold the trolley to help the robot to keep its balance (see section Social Guidance) while it is learning motor commands to produce walking. As we have shown in the section II-A, one of the developmental strategies used to reduce the complexity of the motor learning of a new task is to use motor primitives parameterized by high-level parameters and to tune those parameters rather than motor commands every time.

In this study, motor primitives (see figure 4) are generated by bezier curves which control the angular joints positions over time. Here we generate periodic CPGs using keypoint. CPGs are set by positioning point in the space position/time of each joint. We use seven key-points to allow complex trajectories with multiple change of direction. In order to limit the number of parameter used to describe each key-point, we parameterize only the angular position. The temporal position is set such a way that points be uniformly distributed along the time period of the CPG. We add the last point equal to the first to impose the continuity and then get a periodic signal which can be looped during the simulation. These curves are parameterized by nine high-level parameters:

- 1 parameter describes the period of the curve
- 7 angular parameters uniformely distributed on the entire period parameterize the shape of the bezier curve. They

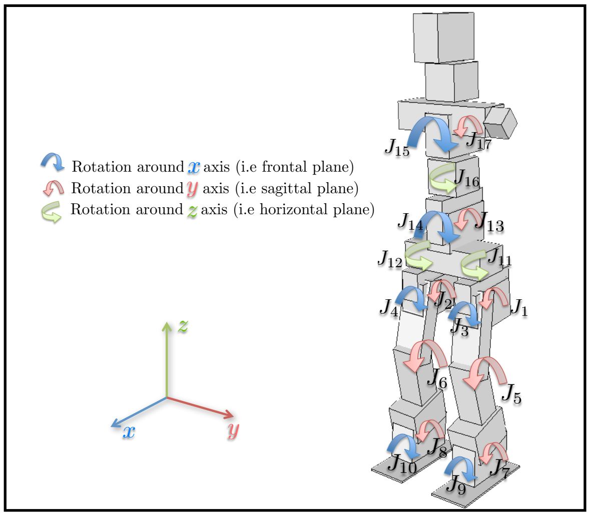


Fig. 3. Overview of actuated joints in our experiments. Names of DoF refer to the order of releasing during learning

are chosen in the explorable space i.e between $\Gamma_{low}$ and $\Gamma_{up}$ (see section IV-A2)

- 1 parameter of phase which allows an entire shifting of the curve (more or less $period/2$) along the time axis

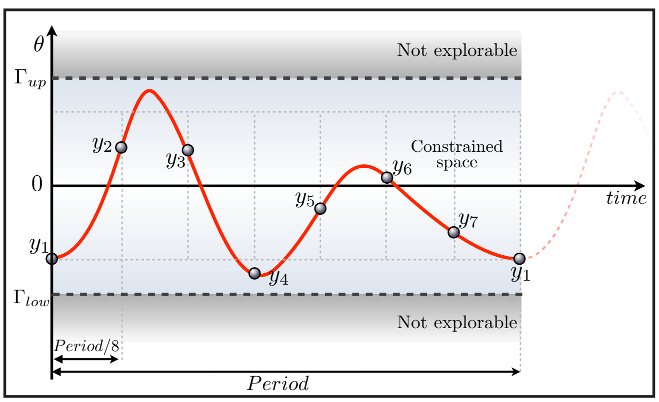


Fig. 4. CPGs are generated by key-points set in the joint position space. This space is constrained (see IV-A2). A bezier curve defined by these key-points generates the motor commands over time.

As shown on the figure 4, an eighth point is added. The last point is equal to the first in order to impose the continuity and then get a periodic signal which can be looped during the simulation. The motor primitive we implemented allows to generate a wide variety of motor control while reducing the dimensions thanks to high-level parameters.

*2) Maturational constraints:* In this paper, our maturational constraints are based on the Bernstein formulation. Indeed, we chosen to gradely release degrees of freedom from the trunk to the tip of the legs. To keep the symmetry of the robot, joints of the right and left legs are released in pairs (see FIgure 3). In our work, a completely constrained joint corresponds to a joint with a motor space equal to zero. Thus, the maturation allows in the same time, the growing of the explorable space for motor primitive and adding new degrees of freedom. Articulations are controlled in position, so with constrain order input. Due to physical or mechanical constraints, the output can be different and be out of bounds but motor commands are choose following maturational constraints.

The global idea is to control all of these constraints using an evolving term $\psi$ called maturational clock. This variable (the x-axis on the Figure 5) represents the ””maturation of brain”” ie the learning evolution. We can set this maturation as a linear function, depending on the time spent in learning. Also, in a developmental robotics frame, we set the maturational clock $\psi$, which controls the evolution of each release of constraint, as depending on the learning activity, and more precisely on competence progress.

Then, we defined a maturational pathway (see figure 5) such as:

$$MotorSpace_{Jointi} = (\psi + (D - i) * tan(\alpha)) * tan(\beta);$$

with parameters :

| | |
|---|---|
| $\alpha$ | angle which defined how many degree of freedom are released in function of $\psi$, |
| $\beta$ | angle which defined how fast the space of each joint is growing according to the maturation $\psi$, |
| $D$ | specifies how many joints are enabled at the beginning of experiment i.e. $\psi = 0$ |
| $i$ | specifies the considered joint. |

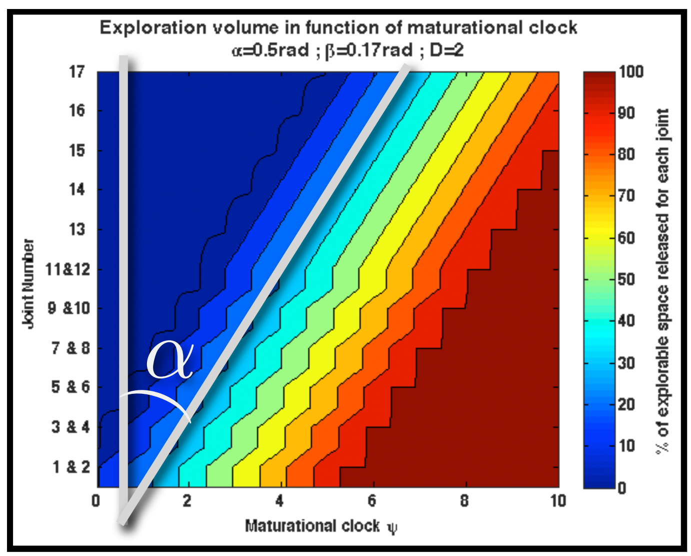


Fig. 5. Releasing way for each degrees of freedom in function of the three parameters ($\alpha$ , $\beta$ and $D$) and the maturational clock $\psi$. With these parameters, at $\psi = 0$, joints 1 & 2 have around 10% of their motor space available. All others joints can not move.

*3) Social guidance: the role of the trolley:* As we described in section II-D. There are some stages, infants did not go directly from crawling to biped walking. They progressively raise their body until be able to walk with balance. Ones of those intermediate stages are hold furniture and pull up to a vertical position, stand while holding onto furniture, take forward steps while holding a caregivers hands, and cruise sideways in an upright position while holding onto furniture. The trolley takes place in one of those stages; indeed, it could be seen as caregivers hand while the robot takes its first forward steps.

In order to reduce the complexity of the learning of walking, we have added a baby trolley to our simulation.The handle of the trolley represent hand of a caregiver and its inertia is sufficient to prevent falling of Acroban. In order to avoid bias on the learning, like the fact that the robot should push the trolley to go forward, the trolley is motorized using a proportional controller pulling slightly the robot in the right direction to guide it (i.e. trying to reach a target velocity). If the gait generated by CPGs is not good enough to follow the trolley, the proportional control reduce the speed until zero or even go back if the robot is falling forward. It is formalize as:

Let us consider the point $W$ CoG of the front wheel and the CoG of the Acroban Pelvis (see Figure 2) . $d(t)$ represent the distance between $W$ and $CoG$ projected along x axis such as $d(t) = ||T(t)CoG(t)||_x$. The trolley controller is defined as:

$$Vel_{trolley}(t) = k * [d(t_0) - d(t)] + V_{target}$$

It could be seen as parents helping their baby to keep his balance. Nevertheless, when motor primitives produce very unbalancing actuator commands, the robot could fall backward, forward or sideward.

### B. Learning algorithm

In this paper we investigate the impact of maturational constraints on the learning efficiency (defined in IV-A2). We evaluate the learning progression in function of the three following parameters (i.e. $\alpha$ , $\beta$ and $D$) in the context of bipedal walking assisted by social guidance. In the next section, we will give more details about the learning method.

*1) Initialization:* In order to initialize the learning, we need to choose a starting set of values parameterizing all motor primitives. We randomly explore the initial space, i.e when $\psi = 0$. In the next paper, we will investigate other starting method based on developmental primitive reflexes.

For each joint, the random motor parameters are chosen among the available range at the initial maturation step ($\psi = 0$). The simulation is launched on each of these 200 sets and we keep the best set, i.e. where the robot was able to traveled the longest distance.

*2) Learning:* The learning algorithm is based on an iterative research of the best solution. For each step we try 30 sets of motor primitives closed to the best set found, then the motor space can increase or not. Here, we release constrained following the gradient of the learning curve. In the next paper, other rules of release will be explored.

**LEARNING LOOP:**

**while** $step \leq maxStep$ **do**

find $k \in \mathcal{N}/fitBest(k) = max(fitBest)$ We keep the global best set as starting point for the next iteration.

- Generates and simulates $n = NlocalExplo$ random new sets $xSet$ among $[\Gamma_{low}, \Gamma_{up}]$

- 60% of these sets are chosen close the best solution $bestSet(k)$ (i.e. more or less 15% on the values).

- 40% are randomly chosen among the available range at the current maturation ($\psi(step)$).

**for** $i = 1 \rightarrow NlocalExplo$ **do**

Simulates walking behavior with the set $xSet(i)$.

$fitNew(i) \leftarrow$ Distance traveled

$timeEnd(i) \leftarrow$Simulation time

Checks if the robot did not fall before the end of simulation :

**if** $timeEnd(i) < MaxSimTime$ **then**

$Fall(i) = true$

**end if**

**end for**

- Keeps only solutions which worked i.e. the robots didn't fall before the end of simulation

- Finds best solution i.e. where the robot was able to traveled the longest distance.

find $m$ such as $fitNew(m) = max(fitNew)$

- Keeps best solution of this step

$fitBest(step) \leftarrow fitNew(m)$

$xBest(step) \leftarrow xNew(m)$

**MATURATION:**

- Evaluates how learning evolved i.e. estimation of the gradient of fitness curve

$d\psi \leftarrow grad(fitBest(step))$

- Evolution of maturational clock

$\psi(step) = max(\psi(step-1) + d\psi; \psi(step-1))$

- If necessary, increases the motor space (i.e. $\psi$ increased)

$[\Gamma_{low}, \Gamma_{up}] \leftarrow MotorSpace_{Jointi=1..17}(\psi, \alpha, \beta, d)$

**end while**

## V. EXPERIMENTS

In this section, we propose to apply the developmental algorithm formalized in the previous section for the learning of walking in a setup involving the Acroban humanoid. The aim of this experiment is to highlight effects of maturational constraints and developmental pathway for the motor learning in high dimensions. Our goal is not to produce the most efficient and faster algorithm to learn how to walk but to evaluate how maturational constraints could help optimization even if we are using a simple optimization algorithm.

### A. Description

We propose to evaluate the effectiveness of learning depending on the parameters of developments in these constraints. In this paper, we evaluate the impact of three parameters on learning. Those parameters $\alpha$ , $\beta$ and $D$ describe the maturational pathway (i.e. how explorable space is growing).

- $\alpha$ describes how fast new joints are added according to the maturation $\psi$,
- $\beta$ describes how fast the space of each joint is growing according to the maturation $\psi$,
- $D$ specifies how many joints are enabled at the beginning of experiment i.e. $\psi = 0$.

With those three parameters, we propose to experiments four cases (see also figure 6):

1) $(\alpha = 0.5, \beta = 0.17, D = 2)$
The initial exploration space is very small; two joints are released and only 10% of their range space is avaible. Then, the space slowly increases, both for the release of new joints or the increase of explorable motor space for each joints.
2) $(\alpha = 0.01, \beta = 1.49, D = 17)$
No constraints, all motor space is available at the beginning of the experiment.
3) $(\alpha = 0.05, \beta = 0.18, D = 17)$
In this case all joints are released from the beginning. But the motor space is small and slowly increase according to the maturation
4) $(\alpha = 0.4, \beta = 0.2, D = 6)$ This is a compromise between 1) and 3). It follows roughly the same law of expansion than 1) but begins with more joints available (i.e. the half of leg articulation)

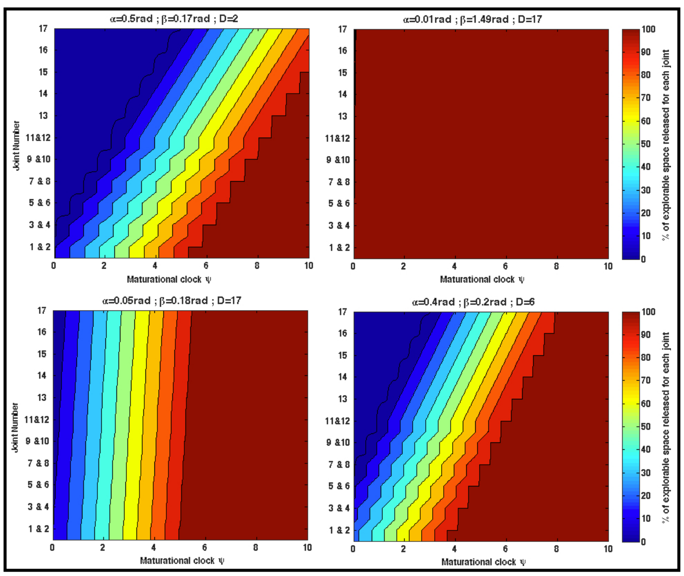


Fig. 6. Figure showing the four parameters sets chosen to compare the learning efficiency in function of the way to release maturational constraints

## B. Results

For each case, we ran several experiments. The figure 7 shows a comparison of the learning success for each case.

As we thought, the experiment 2 completely failed, with our learning algorithm and without any constraints, the robot never managed to find motor primitive producing walking in a short time (only 300 iterations). In addition, around 80% of the trials have led to the robot's falling. So, it is not effective and dangerous for the robot.

Experiment 3 and 4, managed to learn a motor primitive for walking but only at the beginning when motor space was very small. The maturation grown until $\psi = 1$ and after stopped increasing because it couldn't find a better motor primitive.

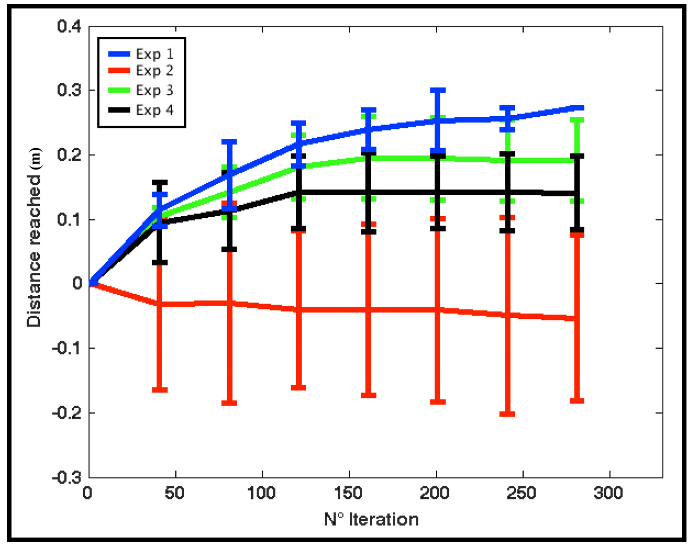


Fig. 7. Learning curves for each experimental parameters. The y-axis represents the distance traveled by the robot and the x-axis the iteration number. The best curve is the blue one with strong maturation constraints. With these parameters, the robot was able to travel around 30cm in 9sec. The best result is close to the velocity of the real robot working with a hand-tuned controller

With a too large motor space, our simple optimization can not find good solutions without many iterations.

On the other hand, the experiment 1 has succeeded the most, reaching 30cm in 9 sec after only 300 iterations. In addition, only 5% of trials failed (i.e. the robot fell). The maturation has grown until $\psi = 4$.The robot is able to explore much more space while taking less risks. Considering in our formalization that the maturation is directly linked to speed of learning, we could conclude that strong maturation constraints lead to a faster and safer way for the learning.

Unfortunately, for technical and time reasons, we could not conduct experiments that lasted over 300 iterations. So we have no information about what happens in the medium term and long-term.

## VI. Discussion

In this paper, we have presented challenges raised by the motor learning in high-dimensions and proposed one way of research: the developmental learning. We proposed a formalization of developmental constraints which can be effective to reduce the complexity of motor learning in high dimensions. We applied the developmental algorithm on a case of learning biped walking with a simulation of our robot Acroban. This first experiment shown better results with maturational constraints similar to the Bernstein's problem. Indeed, the best result correspond to the case where we strongly constrained the motor space and where we released slowly degree of freedom, from the trunk to the tip of legs. Also, this study raises new research axes for the futur work.

We can explore different rules for the evolution of the maturation clock, linearly, depending on the learning or inverse of the learning. We can also explore a new developmental constraint which is to have multiple objective functions to optimize simultaneously. Indeed, learning to walk is not only maximizing the distance reached but also for example minimizing the energy cost, stabilizing the head or walking without trolley. The robot could choose which function it decides to optimize in priority in function of a maturational evolution . For example, trying to go forward and when it succeeds, it tries to go forward without holding the trolley. Then, so far we have used a simple but robust optimization method, it would be meaning to try this kind of algorithm with a more effective method, such as $PI^2$, Natural Actor Critic or Particle Swarm Optimization

Finally, we will do experiments on a real robot, evaluate and compare several learning method with a developmental algorithm based on maturational constraints.

## Acknowledgment

The authors would like to thank all members of the Flowers Team. Many of ideas presented in this paper benefitted from discussions and joint work with many colleagues. The authors also would like to thanks Marc Freeze, founder and development director of v-rep for his fast and effective assistance every time we needed.

This research was partially funded by ERC Starting Grant EXPLORER 240007.

## References

[1] J. Morimoto Nakanishi and al. Learning from demonstration and adaptation of biped locomotion. *Robotics and Autonomous Systems*, 47:79–91, 2004.

[2] Morimoto and Atkeson. Robotics automation magazine. *Robotics and Autonomous Systems*, 2(14):41–51, 2007.

[3] Hirai Hirose and al. The development of the honda humanoid robot. In *Proceedings of the 1998 International Conference on Robotics and Automation*, page 13211326, 1998.

[4] Kaneko Kanehiro and al. Cybernetic human hrp-4c. In *9th IEEE-RAS International Conference on Humanoid Robots. Humanoids 2009*, page 714, 2009.

[5] Oudeyer Baranes, Ly and al. *Developmental constraints on intrinsically motivated skill learning: Towards addressing high-dimensions and unboundedness in the real world*. Baldassarre, Springer, 2010.

[6] W.A. Lee. Neuromotor synergies as a basis for coordinated intentional action. *J. Mot. Behav.*, 16:135–170, 1984.

[7] D'Avella Saltiel and E. Bizzi. Combinations of muscle synergies in the construction of a natural motor behavior. *Nature neuroscience*, 6(3):300–308, march 2003.

[8] L. Ting and J. McKay. Neuromechanics of muscle synergies for posture and movement. *Curr. Opin. Neubiol.*, 7:622–628, 2007.

[9] Cazalets Borde and Clarac. Localization and organization of the central pattern generator for hindlimb locomotion in newborn rat. *The Journal of Neuroscience*, 30(15):49434951, 1995.

[10] DAvella and Portone. Control of fast-reaching movement by muscle synergies combinations. *The Journal of Neuroscience*, 30(26):77917810, 2006.

[11] VR Edgerton, S. Grillner, A. Sjostrom, and P. Zangger. Central generation of locomotion in vertebrates. *Neural control of locomotion*, 18:439–464, 1976.

[12] J. Paillard. Réflexes et régulations d'origine proprioceptive chez l'homme. *Arnette, Paris*, page 293, 1955.

[13] A.J. Ijspeert, A. Crespi, D. Ryczko, and J.M. Cabelguen. From swimming to walking with a salamander robot driven by a spinal cord model. *Science*, 315(5817):1416, 2007.

[14] R. Pfeifer and C. Scheier. *Understanding Intelligence*. MIT Press, Cambridge, MA, 1999.

[15] Rolf Pfeifer and Josh C. Bongard. *How the Body Shapes the Way We Think: A New View of Intelligence (Bradford Books)*. The MIT Press, 2006.

[16] Pfeifer Lungarella and Iida. Self-organization, embodiment, and biologically inspired robotics. *Science*, 318:1088–1093, 2007.

[17] P. Ball. *The self-made tapestry: pattern formation in nature*, volume 212. Oxford University Press New York, 2001.

[18] G. Turkewitz and P. Kenny. Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychology*, 15:357–368, 1982.

[19] L.. Hainline. How the visual system develops: Normal and abnormal development. *Perceptual development: Visual, auditory, and speech perception in infancy*, page 550, 1998.

[20] E. Thelen, L.B. Smith, and Inc NetLibrary. *A dynamic systems approach to the development of cognition and action*. MIT Press, 1994.

[21] Neuroscience and Biobehavioral Reviews. Eye, hand, and trunk control: the foundation for manual development. *Developmental Psychology*, 22:515520, 1998.

[22] L. Berthouze and M. Lungarella. Motor skill acquisition under environmental perturbations: On the necessity of alternate freezing and freeing of degrees of freedom. *Adaptive Behavior*, 12(1):47, 2004.

[23] M. Lungarella and L. Berthouze. Adaptivity through physical immaturity. *Lund University Cognitive Studies*, 2002.

[24] M. Lungarella and L. Berthouze. On the interplay between morphological, neural, and environmental dynamics: a robotic case study. *Adaptive Behavior*, 10(3-4):223, 2002.

[25] B. Vereijken, R.E.A. van Emmerik, HTA Whiting, and K.M. Newell. Free (z) ing degrees of freedom in skill acquisition. *Journal of Motor Behavior*, 24(1):133–142, 1992.

[26] KM Newell and REA Van Emmerik. The acquisition of coordination: preliminary analysis of learning to write. *Human Movement Science*, 8(1):17–32, 1989.

[27] V. Haehl, V. Vardaxis, and B. Ulrich. Learning to cruise: Bernstein's theory applied to skill acquisition during infancy. *Human Movement Science*, 19(5):685–715, 2000.

[28] Ly.O and Oudeyer.PY Lapeyre, M. Bio-inspired vertebral column, compliance and semi-passive dynamics in a lightweight humanoid robot. *IROS 2011 Conference*, 2011.

[29] C. Azevedo, P. Poignet, B. Espiau, B. Brogliato, B. Brogliato, B. Brogliato, B. Brogliato, D. Goeleven, JM Bourgeot, B. Brogliato, et al. Contrôle de la locomotion artificielle: de lhomme aux robots. In *Second MIT Conference on Computational Fluid and Solid Mechanics*, volume 9, pages 322–360, 2003.

[30] J. Rose and J.G. Gamble. *Human walking*. Lippincott Williams & Wilkins, 2006.

[31] C.L. Vaughan. Theories of bipedal walking: an odyssey. *Journal of biomechanics*, 36(4):513–523, 2003.

[32] A. Thorstensson, H. CARLSON, M.R. ZOMLEFER, and J. Nilsson. Lumbar back muscle activity in relation to trunk movements during locomotion in man. *Acta Physiologica Scandinavica*, 116(1):13–20, 1982.

[33] V. Feipel, T. De Mesmaeker, P. Klein, and M. Rooze. Three-dimensional kinematics of the lumbar spine during treadmill walking at different speeds. *European Spine Journal*, 10(1):16–22, 2001.

[34] J.C. Ceccato. *Le tronc, de la locomotion à la commande*. PhD thesis, Universit de Montpellier2, 2009.