Online Adaptation of Locomotion with Evolutionary Algorithms: a Transferability-based Approach

Sylvain Koos
koos@isir.upmc.fr

Jean-Baptiste Mouret
mouret@isir.upmc.fr

ISIR, Université Pierre et Marie Curie-Paris 6, CNRS UMR 7222
4 place Jussieu, F-75252, Paris Cedex 05, France

ABSTRACT

Wheel-legged hybrid robots are versatile machines that can employ several locomotion modes; however, automatically choosing the right locomotion mode is still an open problem in robotics. We here propose that the robot autonomously discovers its locomotion mode using a multi-objective evolutionary optimization and a fixed internal model. Three objectives are optimized: (1) the displacement speed computed with the internal model, (2) the predicted expended energy and (3) the transferability score, which reflects how well the behavior of the real robot is in agreement with the predictions of the internal model. This transferability function is actively learned by conducting 20 experiments on the real robot during the optimization. We tested this approach with a wheel-legged robot in three situations (flat ground, grass-like terrain, tunnel-like environment): in each case, the evolutionary algorithm found efficient controllers for forward locomotion in 1 to 2 minutes.

Categories and Subject Descriptors

I.2.9 [**Computing Methodologies**]: Artificial Intelligence—*Robotics*

General Terms

Algorithms

1. INTRODUCTION

Wheel-legged robots aim at combining the efficiency of wheeled robots with the versatility of legged robots[3, 4, 5]: by adding wheels at the end of legs, they can act like wheeled robots on simple terrains and adapt their posture to the shape of an uneven ground; they can also stop their wheels and be equivalent to a classic legged robot. A few papers investigate different locomotion modes, such as rolling with passive wheels and walking[3]. Nevertheless, none of them tackles one of the most important questions: *how should the robot select its locomotion mode?* And, since there are an infinity of possible situations and an infinity of hybrid locomotion modes, how to discover the best controller in an unforeseen situation? The present paper introduces an evolutionary-based adaptation algorithm to answer these two questions. In the typical scenario, the robot first moves on a flat terrain, in which using the wheels is *a priori* efficient; the robot then encounters a tall grass field in which its wheels are not working anymore: it has to find a new locomotion mode; finally, it leaves the grass field and enters a tunnel with a low ceiling that should never be hit: a new adaptation is required. This scenario illustrates three situations but the goal of the present paper is to introduce a general adaptation algorithm that could be used in any situation and for any robot.

GECCO'11, July 12–16, 2011, Dublin, Ireland.
ACM 978-1-4503-0690-4/11/07.

2. PROPOSED APPROACH

If a robot is in a situation that has never been encountered before, the best thing it can do is to *evolve* a new strategy by itself. Such a situation is a typical use case of reinforcement learning[6] or optimization-based learning[9], but these algorithms require a large number of trials on the real robot, making the learning phase long[1] and potentially dangerous. The Estimation-Exploration Algorithm[1] proposes a faster alternative which relies on an automatically learned, internal dynamic model of the robot: using it, a star-shaped quadruped robot only needed to perform 15 simple actions to learn a new walking behavior after the loss of a leg. However, each time a disagreement is detected between the internal model and the reality, the EEA requires to learn a new internal model from scratch; we think that this complex step is inefficient if the disagreement stems from a change of the environment (most parts of the internal model should still be reliable) and not from a change in the morphology of the robot. Additionally, actions performed by the robot are not goal-directed: the robot can spend a long time to model a part of its morphology which is useless for its goal.

To overcome these limitations, we propose to rely on a similar dynamic internal model, except that it will be provided by the robot's designer. In our approach, the evolutionary algorithm will not modify this model but it will discover what potentially interesting behaviors are not working properly in reality and it will avoid them. To do so, we take inspiration from the transferability approach[7] by formulating the problem as a three-objective optimization process:

$$\text{maximize} \begin{cases} AvgSpeed(x) \\ -Energy(x) \\ Transferability(x) \end{cases}$$

where x is a candidate controller, $AvgSpeed(x)$ is the average displacement speed of the robot in the internal model,

[1]about 3 hours to learn a quadruped locomotion pattern for the Aibo robot[6] and about 20 minutes for a snake-like robot[9].

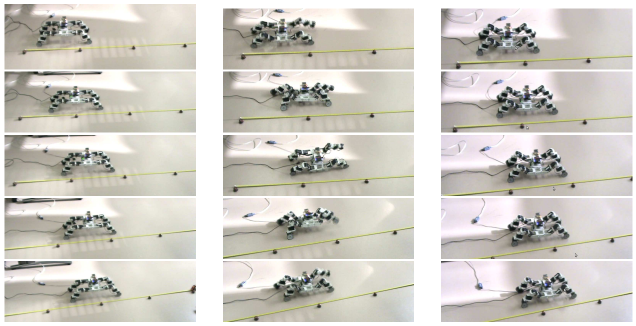

Figure 1: Typical selected behaviors: (left) flat ground: rolling behavior; (center) grass-like: walking behavior; (right) tunnel: rolling behavior with a lowered body.

Table 1: Average speed (cm/s) for each controller and in each situation, standard deviations (5 trials) are in brackets; energy is in arbitrary unit.

tested on / optim. on	flat ground	grass	tunnel	energy
flat ground	17.1 (0.3)	0.0 (0)	8.6 (9.9)	-26 (3)
grass	13.1 (1.9)	13.1 (1.9)	8.3 (7.6)	-33 (1)
tunnel	15.9 (1.7)	0.0 (0)	15.9 (1.7)	-27 (2)

$Energy(x)$ is an estimation of the energy required to perform the behavior (walking uses more energy than rolling) and $Transferability(x)$ is an approximate function which reflects how well the reality matches the prediction of the internal model for the controller x [7]. This last function is approximated by transferring a few well-chosen controllers on the real robot (during the optimization) and subsequently comparing the behaviors in simulation and reality.

The three objectives are optimized simultaneously with NSGA-II, a state-of-the art multi-objective evolutionary algorithm[2]. This optimization algorithm finds an approximate set of all Pareto-optimal trade-offs; to select the final controller, we choose the solution whose distance to the ideal point is minimal among the solutions whose transferability values are higher than a pre-defined threshold[7].

3. EXPERIMENTAL RESULTS

We tested this new approach on a wheel-legged robot inspired by the Hylos robot[4] (figure 1). We investigated three situations to which the robot has to adapt: (1) a flat ground, (2) a grass-like environment (in which wheels are not working at all) and, (3) a tunnel-like environment (in which the maximum height of the robot is constrained). In each case, we allowed 20 tests in reality. The average speed of the robot was evaluated with a motion tracking system but further work will rely on on-board visual odometry[8]. The transferability function was the opposite of the variation of average speeds measured with the internal model and on the robot. The control of the 12-DOFs robot is based on sinusoids, which are parametrized by 3 real numbers: speed of the four wheels, robot's initial posture and movement amplitude of the legs. The goal of the evolutionary algorithm is therefore to find the optimal value of these parameters while taking into account environmental constraints. Energy is crudely approximated by summing the angular movement of each degree of freedom. To obtain statistical results, we performed 5 experiments in each situation.

Results (figure 1) show that: (1) the robot autonomously chose to use its wheels when it was put on a flat ground, (2) it optimized a walking gait when its wheels were unusable and, (3) it lowered its body when it encountered the tunnel. Each optimization phase required 1 to 2 minutes with a recent multi-core computer (including the tests on the robot). Table 1 reports quantitative results: the rolling behavior found on flat ground cannot be used on grass (average speed is null) and is often useless in the tunnel (average speed is low: many controllers did not work), but it is the most energy-efficient controller; the walking mode requires more energy but it appears more versatile; the controllers optimized for the tunnel use slightly more energy than those optimized for flat ground because the robot has to lower its body. *These measures show that adapting the locomotion mode was always useful and often mandatory for the robot to move forward.*

4. REFERENCES

[1] J. Bongard, V. Zykov, and H. Lipson. Resilient Machines Through Continuous Self-Modeling. *Science*, 314(5802):1118–1121, 2006.

[2] K. Deb, S. Agrawal, A. Pratab, and T. Meyarivan. A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II. In *Proc. of PPSN VI*, pages 849–858, 2000.

[3] G. Endo and S. Hirose. Study on roller-walker (multi-mode steering control and self-contained locomotion). In *Proc. of IEEE-ICRA'00*, volume 3, pages 2808–2814. IEEE, 2000.

[4] C. Grand, F. Benamar, F. Plumet, and P. Bidaud. Stability and Traction Optimization of a Reconfigurable Wheel-Legged Robot. *The International Journal of Robotics Research*, 23(10-11), 2004.

[5] P. Jarrault, C. Grand, and P. Bidaud. Large Obstacle Clearance Using Kinematic Reconfigurability for a Rover with an Active Suspension. In *Proc. of CLAWAR'10*, pages 1–8, 2010.

[6] N. Kohl and P. Stone. Policy gradient reinforcement learning for fast quadrupedal locomotion. In *Proc. of IEEE-ICRA'2004*, volume pp, pages 2619–2624, 2004.

[7] S. Koos, J.-B. Mouret, and S. Doncieux. Crossing the Reality Gap in Evolutionary Robotics by Promoting Transferable Controllers. In *Proc. of GECCO'2010*, 2010.

[8] D. Nistér, O. Naroditsky, and J. Bergen. Visual odometry for ground vehicle applications. *Journal of Field Robotics*, 23(1):3–20, 2006.

[9] A. Sproewitz, R. Moeckel, J. Maye, and A. J. Ijspeert. Learning to Move in Modular Robots using Central Pattern Generators and Online Optimization. *The International Journal of Robotics Research*, 27(3-4):423–443, 2008.