

Learning-Based Control for Legged Robotic Locomotion

Julien Moreau

Draft cover. The official EPFL cover is added by the Repro.

Acknowledgements

I thank my thesis director for his trust and support, the members of the jury for their time and questions, and my colleagues in the laboratory for their help with hardware experiments.

Preface

This preface was written by the thesis director. Julien's work bridged two communities of legged robotics that rarely share a controller, and the resulting system has since become a standard baseline in our laboratory.

Abstract

Legged robots can cross terrain that wheels cannot, but controlling them requires reasoning about contacts that make and break many times per second. This thesis combines model-based and learning-based control for quadrupedal locomotion. A reinforcement-learning policy proposes footholds and body motions, and a model-predictive controller turns them into contact forces that respect friction limits. Trained entirely in simulation and deployed without fine-tuning, the controller walks over rough terrain, steps and slopes, and reduces falls on a rubble course from 14 to 2 in 100 trials compared with a purely learned policy.

Keywords: legged robots, reinforcement learning, model-predictive control, sim-to-real transfer

Résumé

Les robots à pattes franchissent des terrains inaccessibles aux roues, mais leur commande exige de gérer des contacts qui s'établissent et se rompent plusieurs fois par seconde. Cette thèse combine une politique apprise par renforcement, qui propose les appuis et le mouvement du corps, et une commande prédictive qui en déduit des forces de contact respectant les limites de frottement. Entraîné en simulation et déployé sans ajustement, le contrôleur réduit les chutes sur un parcours d'éboulis de 14 à 2 sur 100 essais par rapport à une politique purement apprise.

Mots-clés : robots à pattes, apprentissage par renforcement, commande prédictive, transfert de la simulation au réel

Contents

Acknowledgements	i
Preface	ii
Abstract	iii
Résumé	iv
1 Introduction	1
2 Background	2
3 A Hierarchical Controller	3
4 Hardware Experiments	4
5 Conclusion	5
A Simulation Randomization	6
Curriculum Vitae	8

List of Figures

3.1	The learned policy plans; the predictive controller enforces physics.	3
-----	---	---

List of Tables

4.1	Falls in 100 trials per course.	4
-----	---	---

Chapter 1

Introduction

Walking robots have left the laboratory, but uneven ground, slippery surfaces and unexpected obstacles still cause falls. Model-based controllers give guarantees about forces and friction but rely on accurate terrain models; learned policies are robust to modelling errors but offer no guarantees [2]. This thesis lets each approach do what it does best.

Chapter 2

Background

For a quadruped with stance feet at positions r_i applying ground forces f_i , the centroidal dynamics are

$$m\ddot{c} = \sum_i f_i + mg, \quad \dot{L} = \sum_i (r_i - c) \times f_i, \quad (2.1)$$

with each force constrained to its friction cone. Model-predictive control optimizes these forces over a short horizon [1]; reinforcement learning trains policies by trial and error in simulation [3].

Chapter 3

A Hierarchical Controller

A policy network observes proprioception and a local height map and outputs footholds and a body trajectory at 50 Hz. A model-predictive controller at 500 Hz computes contact forces that track the trajectory while satisfying Equation (2.1) and the friction constraints.

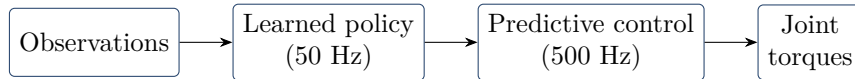


Figure 3.1: The learned policy plans; the predictive controller enforces physics.

Chapter 4

Hardware Experiments

Table 4.1 reports falls over 100 trials on three courses.

Table 4.1: Falls in 100 trials per course.

Controller	Steps	Slopes	Rubble
Model-based only	9	4	21
Learned only	3	2	14
Hierarchical	1	0	2

Chapter 5

Conclusion

Splitting locomotion into a learned planner and a model-based force controller combines the robustness of learning with the guarantees of optimization. Adding vision-based terrain perception to the policy is the next step.

Appendix A

Simulation Randomization

During training, friction coefficients were sampled from 0.3 to 1.2, payloads from 0 to 5 kg, and motor strengths from 80% to 110% of nominal.

Bibliography

- [1] Jared Di Carlo, Patrick M. Wensing, Benjamin Katz, Gerardo Blede, and Sangbae Kim. Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1–9, 2018.
- [2] Jemin Hwangbo, Joonho Lee, Alexey Dosovitskiy, Dario Bellicoso, Vassilios Tsounis, Vladlen Koltun, and Marco Hutter. Learning agile and dynamic motor skills for legged robots. *Science Robotics*, 4(26):eaau5872, 2019.
- [3] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv:1707.06347, 2017.

Curriculum Vitae

Julien Moreau

2021 to 2026 Doctoral assistant, EPFL, Lausanne

2019 to 2021 M.Sc. in Robotics, Vertex Institute of Technology

2016 to 2019 B.Sc. in Mechanical Engineering, Vertex Institute of Technology