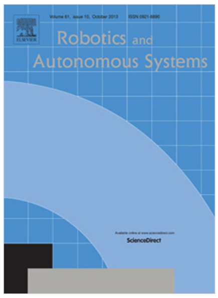
Accepted Manuscript



Benchmarking model-free and model-based optimal control

Ivan Koryakovskiy, Manuel Kudruss, Robert Babuška, Wouter Caarls, Christian Kirches, Katja Mombaur, Johannes P. Schlöder, Heike Vallery

PII: S0921-8890(16)30159-2
DOI: http://dx.doi.org/10.1016/j.robot.2017.02.006
Reference: ROBOT 2797

To appear in: *Robotics and Autonomous Systems*

Please cite this article as: I. Koryakovskiy, et al., Benchmarking model-free and model-based optimal control, *Robotics and Autonomous Systems* (2017), http://dx.doi.org/10.1016/j.robot.2017.02.006

Benchmarking model-free and model-based optimal control☆

Ivan Koryakovskiy[a], Manuel Kudruss[*,b], Robert Babuška[e], Wouter Caarls[c,d], Christian Kirches[b], Katja Mombaur[b], Johannes P. Schlöder[b], Heike Vallery[a]

[a]Delft Biorobotics Lab, Faculty of Mechanical Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands
[b]Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany
[c]Department of Computer Science, Federal University of Rio de Janeiro, 22451-900 Rio de Janeiro, RJ Brazil
[d]Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, 21941-909 Rio de Janeiro, RJ Brazil
[e]Delft Center for Systems and Control, Faculty of Mechanical Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands

Abstract

Model-free reinforcement learning and nonlinear model predictive control are two different approaches for controlling a dynamic system in an optimal way according to a prescribed cost function. Reinforcement learning acquires a control policy through exploratory interaction with the system, while nonlinear model predictive control exploits an explicitly given mathematical model of the system. In this article, we provide a comprehensive comparison of the performance of reinforcement learning and nonlinear model predictive control for an ideal system as well as for a system with parametric and structural uncertainties. The comparison is based on two different criteria, namely the similarity of trajectories and the resulting rewards. The evaluation of both methods is performed on a standard benchmark problem: a cart-pendulum swing-up and balance task. We first find suitable mathematical formulations and discuss the effect of the differences in the problem formulations. Then, we investigate the robustness of reinforcement learning and nonlinear model predictive control against uncertainties. The results demonstrate that nonlinear model predictive control has advantages over reinforcement learning if uncertainties can be eliminated through identification of the system parameters. Otherwise, there exists a break-even point after which model-free reinforcement learning performs better than nonlinear model predictive control with an inaccurate model. These findings suggest that benefits can be obtained by combining these methods for real systems being subject to such uncertainties. In the future, we plan to develop a hybrid controller and evaluate its performance on a real seven-degree-of-freedom walking robot.

Key words: Reinforcement Learning, Optimal Control, Nonlinear Model Predictive Control, Parametric Uncertainties, Structural Uncertainties

1. Introduction

In robotics, one cannot expect to work with ideal models of the systems under control, or of their environments. Rather, we have to face unforeseen situations and unknown conditions, and aim for reactions that are feasible and, ideally, optimal with respect to given task performance criteria. A typical task is bipedal locomotion, where a robot needs to maintain stability and pace on an uneven floor with uncertain roughness and slope [1].

Two common approaches to control dynamic systems are Nonlinear Model Predictive Control (NMPC) and Reinforcement Learning (RL). Both approaches can cope with uncertainties in the form of model-plant mismatch. Reinforcement learning has been proven suitable as a real-time closed-loop control concept in robotics [2], and NMPC in industry [3]. However, the use of NMPC in robotic applications, especially humanoid robotics and bipedal walking, is still an open research field [4, 5, 6].

In this article, we use a swing-up and balancing problem for a *cart-pendulum* system [7, 8] to quantitatively assess both control approaches. Our choice of this benchmark problem is motivated by the fact that main features of passive dynamic walking can be modeled by an inverted pendulum [9]. The same equivalence holds for the upper body of a more detailed model of a bipedal walker. The study presented in this article highlights the differences in performance of NMPC and RL under structural and parametric uncertainties for this benchmark problem.

Nonlinear model predictive control. Nonlinear model predictive control is a closed-loop control strategy in which the control action at the current sampling instant is computed by solving an open-loop optimal control problem over a finite prediction horizon. NMPC, as a model-based optimal control method, relies on a given mathematical model of the real-world system to be controlled. In this context, advanced direct methods of optimal control, see the survey [10], are the methods of choice for computing NMPC feedback control actions in real-time.

For NMPC, full state and parameter information of the model is required to compute the control action. Whenever the full state is not measurable or model parameters are not exactly known, methods of on-line state and parameter estimation have

☆The first two authors contributed equally to this work.
*Corresponding author