

On the Role of Models in Learning Control

Citation for published version (APA):

Poot, M. M., Portegies, J., & Oomen, T. (2020). On the Role of Models in Learning Control: Actor-Critic Iterative Learning Control. In R. Carloni, B. Jayawardhana, & E. Lefeber (Eds.), *39th Benelux Meeting on Systems and Control: Book of Abstracts* (pp. 155). University of Groningen.
<https://www.beneluxmeeting.nl/2020/uploads/papers/boa.pdf>

Document status and date:

Published: 01/01/2020

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

On the Role of Models in Learning Control: Actor-Critic Iterative Learning Control

Maurice Poot^{1,*}, Jim Portegies², Tom Oomen¹

¹Control Systems Technology Group, Dept. of Mechanical Engineering, Eindhoven University of Technology, The Netherlands

²CASA, Dept. of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands. *Email: m.m.poot@tue.nl

1 Background

Learning has large benefits for control applications, including high-tech mechatronic systems, by greatly improving the accuracy using data from past tasks. Norm-Optimal Iterative Learning Control (NOILC) exploits model knowledge for fast and safe learning [1]. Obtaining a model of a system leads to undesired user-intervention. In reinforcement learning (RL), many model-free learning techniques are developed that show promising convergence properties [2, 3].

2 Problem formulation

Although ILC methods often have fast and safe convergence properties and exceptional performance, these methods require an explicit system model. The aim of this research is to investigate model-based and model-free learning for mechatronic systems from the perspective of prior model-knowledge and sample complexity and to develop a model-free approach to learn the optimal feedforward signal from experiment data.

3 Approach

A model-free approach to learn the optimal feedforward signal is developed and is called actor-critic iterative learning control (ACILC). ACILC exploits the use of feedforward parameterization with basis functions such that implicit model knowledge can be incorporated and uses the actor-critic algorithm of RL [2, 3] to learn the feedforward parameters without explicitly using a system model.

4 Results

Initial results for positioning of a consumer printer demonstrate the model-free performance of ACILC in comparison to NOILC [1] with basis functions that uses an explicit system model. In Figure 1, the cost per trial is shown, demonstrating that ACILC achieves the same optimal cost as NOILC in 20 trials for the same basis functions. The cost of ACILC varies significantly in the first trials due to exploration necessary for learning, while the cost of the NOILC method converges in one step due to the use of a model.

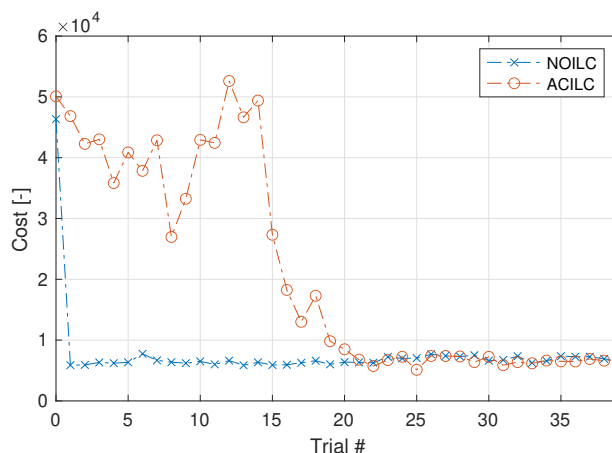


Figure 1: Cost per trial for model-based NOILC with basis functions and model-free ACILC experiments.

5 Conclusion and outlook

The developed ACILC framework is a model-free framework capable of learning the optimal feedforward signal with little implicit model knowledge incorporated in the basis functions. Future research focuses on tuning of the actor-critic parameters and extending the basis functions to further improve the convergence performance.

6 Acknowledgements

This work is supported by ASM Pacific Technology and NWO-VIDI nr 15698.

References

- [1] Oomen, T. (2018). Learning in machines: towards intelligent mechatronic systems through iterative control. *Mikroniek*, 6, 5-11.
- [2] Grondman, I., Buşoniu, L., Lopes, G., and Babuška, R. (2012). A survey of actor-critic reinforcement learning: standard and natural policy gradients. *IEEE Trans. on Systems, Man, and Cybernetics, Part C (Appl. and Rev.)*, 42(6), 1291-1307.
- [3] Recht, B. (2019). A tour of reinforcement learning: the view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2, 253-279.