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

Maurice Poot [1,*], Jim Portegies [2], Tom Oomen [1]
[1]Control Systems Technology Group, Dept. of Mechanical Engineering, Eindhoven University of Technology, The Netherlands
[2]CASA, 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.

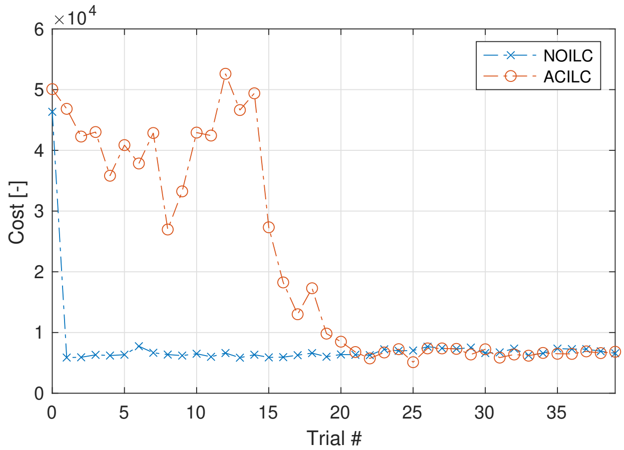

Figure 1: Cost per trial for model-based NOILC with basis functions and model-free ACILC experiments.

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.