Continuous State Space Q-Learning for control of Nonlinear Systems

ten Hagen, S.H.G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Summary

The topic in this thesis is the use of Reinforcement Learning (RL) for the control of real systems. In RL the controller is optimized based on a scalar evaluation called the reinforcement. For systems with discrete states and actions there is a solid theoretic base and convergence to the optimal controller can be proven. Real systems often have continuous states and control actions. For these systems, the consequences of applying RL is less clear. To enhance the applicability of RL for these systems, more understanding is required.

One problem when RL is applied to real continuous control tasks is that it is no longer possible to guarantee that the closed loop remains stable throughout the learning process. This makes it a dangerous way to obtain a controller, especially since a random process called “exploration” has to be included during training. Another problem is that most RL algorithms train very slowly. This means that for a slow process the application of RL is very time consuming.

When the system is linear and the costs are given by a quadratic function of the state and control action, Linear Quadratic Regularization (LQR) can be applied. This leads to the optimal linear feedback function. System Identification (SI) can be used in case the system is unknown. The LQR task can also be solved by Q-Learning, a model-free RL approach in which the system is not explicitly modeled, but its cost function is. We called this method LQRQL. To find the optimal feedback it is necessary that sufficient exploration is used. We expressed the performance of the resulting feedback as a function of the amount of exploration and noise. Based on this we derived that a guaranteed improvement of the performance requires that more exploration is used than the amount of noise in the system. Also we derived that the LQRQL approach requires more exploration than the SI approach.

Most practical systems are nonlinear systems, for which nonlinear feedback functions are required. Existing techniques for nonlinear systems are often based on local linear approximations. The linear feedbacks of the LQRQL and SI approach are not always able to give a good local linear approximation of a nonlinear feedback function. It is possible to extend the LQRQL approach. The Extended LQRQL approach estimates more parameters and results in a linear feedback plus a constant. In an experiment on a nonlinear system we showed that the extended LQRQL approach gives a better local approximation of the optimal nonlinear feedback function.

For nonlinear systems, function approximators can be used to approximate the Q-function. We showed that it is possible to combine the LQRQL approach with a feedforward neural network approximation of the Q-function. The nonlinear feedback function can directly be determined from the approximated Q-function, by first computing a linear feedback for which a nonlinear correction can be determined. This results in a global nonlinear feedback function. Experiments have shown that this approach requires less training data than an approach based on partitioning of the state space, where each partition approximates a linear feedback.

Reinforcement Learning can be used as a method to find a controller for real systems. To an unknown linear system the LQRQL approach can be applied. The Neural Q-Learning approach can be used when the system is nonlinear.