Abstract - Simulation-Based Optimization with Non-Parametric State Space Approximation using Reinforcement Learning

Name of the Candidate: Daniel Pasterk
Name of the Supervisors: Felix Breitenecker, Andreas Körner
Name of the 1st Referee: Oliver Rose
Name of the 2nd Referee: Thorsten Pawletta

Decision processes play a major role in a wide range of applications. The stochastic nature and inherent complexity of these processes can, in certain cases, make an analytical solution difficult or impossible. Simulation-based optimization, together with machine learning methods, represents a feasible solution for these cases.

In the proposed approach, optimization problems are modeled as Markov decision processes and solved using reinforcement learning. For large and complex problems, the state space has to be approximated by a suitable method. Current approaches based on artificial neural networks enable powerful algorithms, but they often are associated with stability problems. In addition, performance is highly correlated with the correct choice of hyperparameters, which are problem-dependent. This makes it difficult to use reinforcement learning as an optimization method for yet unknown problems.

This thesis addresses the need for reliable optimization methods for discrete-event models. It develops algorithmic approaches based on ideas and concepts from non-parametric methods. While parametric models often make a specific assumption about the underlying data distribution, non-parametric methods do not. This allows them to adapt flexibly to the actual structure of the data.
Discrete-event simulation models can have a state space whose structure is very irregular. The actual size and complexity are a priori unclear. It turns out that these models can benefit greatly from the chosen approach and also provide the necessary amount of data to extract optimal solution strategies. The research questions address the applicability of these non-parametric methods as a solution method for the underlying Markov decision processes.

The applicability is demonstrated on several stochastically challenging case studies that are highly relevant in logistics, medicine and operational research. The possible performance of the algorithms, as well as their methodological limitations, are quantified with simulations. The methods prove to be very powerful when combined with a good modeling of the problem. The quality of the solution strategies reaches a high level in almost all of the ten very diverse problem definitions and often outperforms Deep Reinforcement Learning (DQN). The extracted decisions remain comprehensible and are also suitable for risk-sensitive problems where artificial neural networks show a problematic lack of transparency. Furthermore, integration into the framework of policy gradient methods is possible, where the presented state-space approximations were used as a stable baseline. This enables very powerful algorithms based on the Actor-Critic architecture.

This work highlights the potential of non-parametric methods in reinforcement learning, offering a viable alternative to traditional deep learning methods for simulation-based optimization.