Bram de Moor

The use of known heuristics to improve the performance of deep reinforcement learning in inventory management.

Co-authors: Joren Gijsbrechts, Robert Boute

Deep reinforcement learning (DRL) has proven to be an effective, general-purpose technology to develop 'good' replenishment policies in inventory management. We show how transfer learning from existing, well-performing heuristics may stabilize the training process and improve the performance of DRL in inventory control. While the idea is general, we specifically implement potential-based reward shaping to a deep Q-network algorithm to manage inventory of perishable goods that, cursed by dimensionality, has proven to be notoriously complex. The application of our approach may not only improve inventory cost performance and reduce computational effort, the increased training stability may also help to gain trust in the policies obtained by black box DRL algorithms.