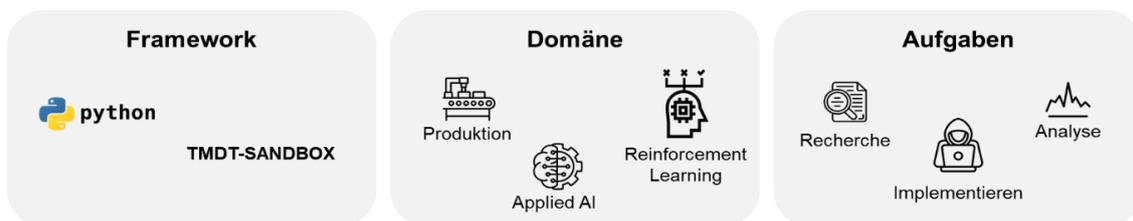
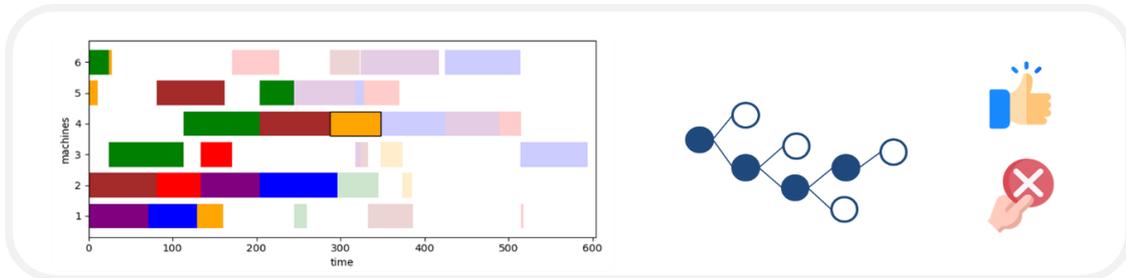


## Ausschreibung Masterarbeit

# Nutzung von Monte-Carlo Search für Reinforcement Learning Basierte Produktionsplanung



### Ausgangslage

Effektive Produktionsplanung ist eine wichtige Voraussetzung produzierender Unternehmen, um sich im internationalen Wettbewerb behaupten zu können. Teil davon ist zu entscheiden, wann welche Teile auf welchen Maschinen bearbeitet werden müssen. Für komplexe Produktionssysteme ist diese Entscheidungsfindung ein ungelöstes Problem. In den letzten Jahren ist durch immer potentere Verfahren des maschinellen Lernens, insbesondere dem Reinforcement Learning, eine neue aussichtsreiche Lösungsmöglichkeit hinzugekommen. Hierbei werden gute Aktionssequenzen autonom durch Interaktion mit einer Simulationsumgebung auf Basis von Belohnung und Bestrafung erlernt.

### Problemstellung

Die Belohnung und Bestrafung einzelner Aktionen stellt bei Produktionsplanungen eine besondere Herausforderung dar, da sich die Folgen einzelner Aktionen oft erst später herausstellen. Gesucht wird daher nach einer Möglichkeit, den Einfluss einzelner Aktionen auf das zukünftige Gesamtergebnis besser abschätzen zu können. Eine Möglichkeit dafür ist die teilweise Vorhersage/Simulation der Zukunft über Monte-Carlo Simulationen.

### Vorgehensweise und Erwartete Ergebnisse

In der Arbeit sollen Möglichkeiten der Verwendung einer Monte-Carlo Simulation recherchiert, in Python implementiert und analysiert werden. Auf Basis der Recherche wird ein bestehendes RL- und Simulations-Framework insbesondere durch das Weiterspielen von Aktionssequenzen und eine nutzbringende Bewertung dieser Sequenzen erweitert. Die Ergebnisse werden mit dem existierenden Ansatz ohne Monte-Carlo Simulation verglichen.

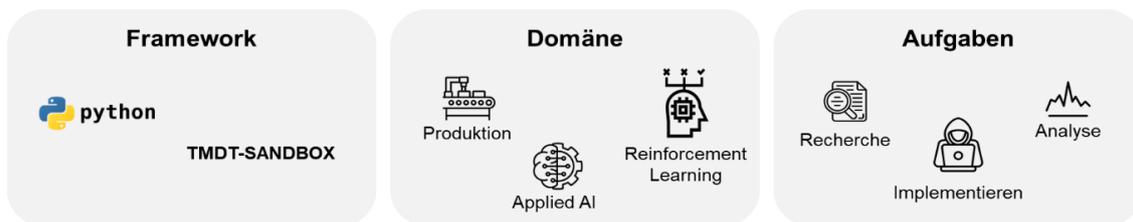
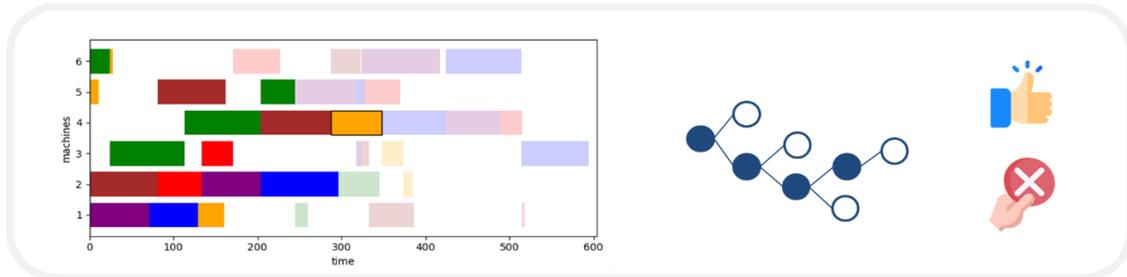
### Ansprechpartner

Constantin Waubert de Puiseau

Tel.: +49 202 439 1713 | E-Mail: [waubert@uni-wuppertal.de](mailto:waubert@uni-wuppertal.de)

Master's Thesis Topic

# Leveraging Monte-Carlo Search for Reinforcement Learning Based Production Scheduling



## Initial Situation

Effective production planning is an important prerequisite for manufacturing companies to be able stay internationally competitive. Part of this is to decide when which parts have to be processed on which machines. For complex production systems this decision making remains a challenging problem. In recent years, increasingly potent machine learning methods, especially reinforcement learning (RL), have added a new promising solution possibility. Here, good action sequences are learned autonomously by interacting with a simulation environment based on reward and punishment.

## Problem Definition

Rewarding and punishing individual actions is a particular challenge in production planning, since the consequences of individual actions often only become apparent later. Therefore, a possibility is sought to better estimate the influence of individual actions on the future overall result. One possibility for this is the partial prediction of the future via Monte-Carlo simulations.

## Methods and Expected Results

In the thesis, possibilities of using a Monte-Carlo simulation will be researched, implemented and analyzed. On the basis of the research, an existing RL and production planning framework will be extended, in particular by playing out sequences of actions and a useful evaluation of these sequences. The results are compared with the existing approach without Monte-Carlo simulation.

## Contact Person

Constantin Waubert de Puiseau

Tel.: +49 202 439 1713 | E-Mail: [waubert@uni-wuppertal.de](mailto:waubert@uni-wuppertal.de)