EPSRC CDT in Distributed Algorithms

PhD Project: Bayesian Reinforcement Learning for Control of Continuous Industrial Processes

University of Liverpool

PhD Student: Oliver Dippel
Project Partner: NSG

Supervisors:
Dr. Bei Peng, University of Liverpool
Dr. Alexei Lisitsa, University of Liverpool
Ian Williams, NSG

Project Description

Reinforcement Learning (RL) describes a goal-directed learning approach from interactions. Commonly one tries to maximise an overall value, expressed by cumulative future discounted rewards. The overall value can be represented by many things such as winning a game or minimizing a cost function. The reward itself in each state represents a payoff realised by an action taken in that state. Both the state and the possible actions and their corresponding reward can be manifold. States can, for instance, describe the game situation of a chess game. A state is then any arrangement of the chess pieces on the board. Given this situation, the "agent" can now choose from the abundance of his actions to rearrange one of his pieces on the board. The reward can then be described, for example, by the removal of an opponent's piece. The aim is to "learn" a policy that wins the game or, in other words, maximises the total reward. The Bayesian part allows the additional consideration of maintaining a distribution over various quantities such as the parameters, the value function, policy or its gradient.

This project is focussed on finding a good control strategy to optimize the quality of a glass product during a specific manufacturing process or reducing the cost of this process through less faulty goods. Ideally, both goals may be combined. The widely used Proportional-Integral-Differential (PID) controller should thus be replaced or expanded at best. Although the PID controller is widely used, it has its weaknesses and sometimes requires human intervention. To overcome them Bayesian Inference will be used to generate simulated data based on offline historic data from the manufacturing process. Then the simulated data serves as an environment for the RL approach, which can learn how to best apply the control automatically. Optimally, not only offline data is used later, but also online data, with which the RL algorithm serves as a controller in the active production process. And thereby either reduces the costs of the production process, increases the quality of the glass or a mixture of both.

Go to the EPSRC CDT In Distributed Algorithms website.