



Contribution ID: 61

Type: **not specified**

## **Optimising for average reward in a continuing environment: an application to industrial production planning**

*Tuesday, 17 September 2024 11:30 (30 minutes)*

Our research addresses the industrial challenge of minimising production costs in an undiscounted, continuing, partially observable setting. We argue that existing state-of-the-art reinforcement learning algorithms are unsuitable for this context. We introduce Clipped Horizon Average Reward (CHAR), a method tailored for undiscounted optimisation. CHAR is an extension applicable to any off-policy reinforcement learning algorithm which exploits known characteristic times of environments to simplify the problem. We apply CHAR to an industrial gas supplier case study and demonstrate its superior performance in the specific studied environment. Finally, we benchmark our results against the standard industry algorithm, presenting the merits and drawbacks of our approach.

### **Type of presentation**

Talk

### **Classification**

Mainly application

### **Keywords**

reinforcement learning, production planning, industrial application

**Primary author:** BERHAUT, Paul (Air Liquide)

**Co-author:** Mr SAMPAIO DOS SANTOS BRANDAO, Patrick (Air Liquide)

**Presenter:** BERHAUT, Paul (Air Liquide)

**Session Classification:** frEnbis invited session: Deep learning in industry

**Track Classification:** AI in Industry