Available online at www.sciencedirect.com
ScienceDirect
Procedia CIRP 51 (2016) 57 – 62
www.elsevier.com/locate/procedia

3rd International Conference on Ramp-up Management (ICRM)

# Automated Production Ramp-up Through Self-Learning Systems

Philipp Ennen*, Sebastian Reuter, Rene Vossen, Sabina Jeschke

*Institute of Information Management in Mechanical Engineering (IMA)*

* Corresponding author. Tel.: +49-241-80-91192 ; fax: +49-241-80-91122. *E-mail address:* philipp.ennen@ima-zlw-ifu.rwth-aachen.de

**Abstract**

The ramp-up of production systems is characterised by situations that arise for the first time. Due to the unpredictability of system behaviour in such situations, instabilities occur that lead to reduced production effectiveness. In order to deal with the resulting uncertainty, this paper presents an approach for self-directed systems capable of "learning", that is, they adapt their behaviour depending on the signals and changes of the circumfluent world. The advantages of such systems are significant, as they can react to changing products, production equipment and process constraints, and are able to function in exceptional situations. The presented concept makes use of reinforcement learning, one of the most general approaches to learning control. Simulations of three different ramp-up processes are used, where, as a demonstration, robots have to assemble windscreens on a moving truck.

Peer-review under responsibility of the scientific committee of the 3rd International Conference on Ramp-up Management (ICRM)
*Keywords:* artificial intelligence; cyber physical production systems; ramp-up; reinforcement learning; automation; robotics

## 1. Introduction

Against the backdrop of shrinking product lifecycles and the desire for customised products, industrial enterprise aims to cut the time taken to reach full capacity utilisation in production. This requires multiple and fast production ramp-ups as products change more often. Thereby the *production ramp-up* describes the phase between the end of the product development and the full capacity production [1]. During ramp-up, the production system's behaviour is not predictable as situations are arising for the first time [2]. In automation, this leads to exceptional situations that disrupt the ramp-up and require an adaption of the system. Therefore, flexible and easily adaptive automation is necessary [3]. Cyber-physical production systems (CPPS) are a possible solution for future manufacturing systems [4]. They have the capability to interact with the physical environment as well as the digital world [5], be aware of its ideal performance, be able to measure its current performance against that of the ideal, and be able to adapt its actions for improving its behaviour [6]. However, such CPPS needs a concept for self-learning as it enables related subsystems – referred to as CPPS-agents – to adapt to unknown situations on their own.

**Nomenclature**

| | |
|---|---|
| $\pi$, q | Policy |
| $y, \dot{y}$ | State variables |
| $\ddot{y}$ | Action variables |
| $x$ | State of canonical system |
| $\alpha_y, \beta_y, g$ | Parameters of the damped linear system |
| $\boldsymbol{\theta}$ | Adaptable/learning parameters of a policy |
| $\tau$ | Scaling factor |
| $\boldsymbol{w}$ | Weight vector |
| $f(x)$ | Forcing function |
| $\Psi_i(x)$ | Gaussian basis function |
| $\varepsilon$ | Relative entropy |
| $R(\boldsymbol{\theta})$ | Reward for parameters $\boldsymbol{\theta}$ |
| $p_{TCP}, p_r$ | Positions |
| $t$ | time step of the system |

This paper introduces a general approach based on findings within the field of artificial intelligence and reinforcement learning in particular. The proposed CPPS-agents do not require any prior knowledge of their mechanical structure or of the environment the machine is active in. Instead, they have the ability to immediately learn an end-to-end (sensors to actors)


Peer-review under responsibility of the scientific committee of the 3rd International Conference on Ramp-up Management (ICRM)
doi:10.1016/j.procir.2016.05.094

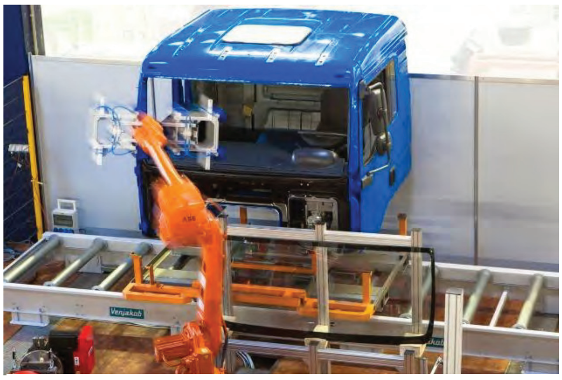

Fig 1 Assembly under flow-conditions

control strategy from trial and error interactions with the circumfluent world. These will be evaluated and subsequently integrated into the control strategy for future behaviour.

In the following chapters, we introduce the general architecture of a self-learning CPPS-agent, then we present the state of the art aspects of reinforcement learning and subsequently present an approach to self-learning for a CPPS-agent. Finally, we show the potential of such an approach within several simulated ramp-up processes orientated to assembly systems under flow-conditions (cf. Figure 1).

Flow assembly could reduce assembly time by as much as 60% [7], though it requires compensation for the dynamic influences of the continuously moving parts, especially in the case of large-scale parts. In the scenario, a robot has to assemble a windscreen on a moving truck using an external reference system – the iGPS (Nikon Metrology) – to sense the position of both the robot and the truck. Due to a nonstationary setup, the assembly process requires the synchronisation of the robot and the moving truck. This gets more difficult as the truck begins to oscillate in motion. Furthermore, the setup is characterised by a long tolerance chain, composed of tolerances in the mechanical structure of the robot, the perception, the action execution and the product. Therefore, pre-programming a synchronised movement is difficult to achieve. An online adaption of the process is also difficult as the external reference system has a high latency time (~ 1 second). However, using the concept of a self-learning CPPS-agent, the robot learns an end-to-end control strategy that enables a synchronised movement. Thereby, all related dynamics (i.e. the moving truck) and the long tolerance chain are considered, which are assumed to be recurring for the purpose of this paper.

### *1.1. Architecture for self-learning CPPS-agents*

Much research has been done on model-based approaches that control an agent's behaviour. This approach reduces the search space of an agents' behaviour by integrating prior knowledge into the decision making process. As an example, autonomous robots plan their behaviour based on a word-model by taking the knowledge of uncertainty in sensing and acting into account [8]. However, during ramp-up, this knowledge is not available as the situations are arising for the first time. In contrast, self-learning CPPS-agents are able to deal with such unknown situations. They optimise their behaviour regarding external feedback without the need for models.

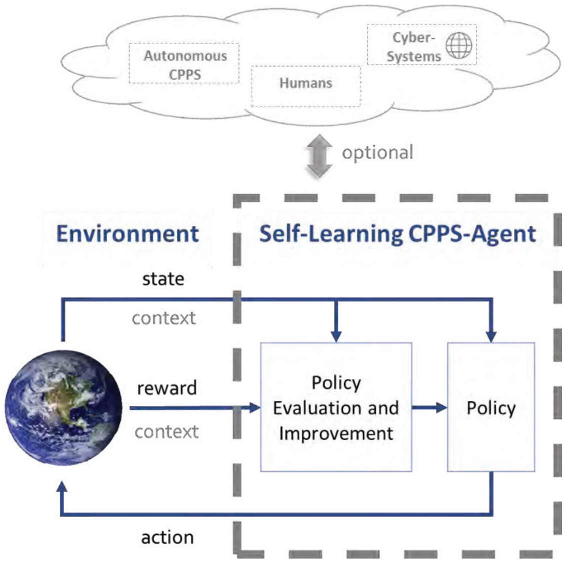


Fig 2 Self-learning CPPS-agent

As pointed out in Figure 2, a self-learning CPPS-agent requires three core components:

- Policy
- Policy evaluation and improvement
- External reward

The *policy* is the decision function of an agent's behaviour, which maps an observed state to an action. This process can be discrete or continuous, therefore the Markov-assumption, in which the observed state space represents the whole history of the CPPS-agent, is consulted [9]. A self-learning CPPS-agent should be capable of producing a broad range of products. Therefore, the process is carried out hierarchically and switches to a suitable behaviour fitting the context. The context is either directly communicated or estimated based on the current state, which allows the process to be used in the production of multi-variant products in particular. The *policy evaluation* checks the current performance of the policy with respect to an *external reward,* whereas the *policy improvement* optimises the policy with the objective of maximising the estimated future rewards. The rewards are assigned by entities within the environment, like the product, another CPPS-agent or a human for instance. However, within ramp-up, a CPPS-agent never knows what the reward will be exactly, so policy improvement is solely based on the estimation explained later.

This so-called reinforcement-learning framework is a naturally inspired concept that facilitates broad use across all systems, whilst having the capability to act and observe [10]. This approach operates without any reliable communication channels to humans or other systems. While the policy improvement can be supported by information provided by other (cyber-) systems, the self-learning CPPS-agent does not necessarily require this information.