

DOCTORAL (PhD) THESES

---

**Reinforcement learning  
applications for selected  
industrial optimization problems**

---

*Author:*

Tamás KEGYES

*Supervisors:*

Dr. Zoltán SÜLE

Prof. Dr. habil. János ABONYI

University of Pannonia

Faculty of Information Technology

Doctoral School of Information Science and Technology

2025

# 1 Motivations and objectives

My research aims to develop new methods and algorithms and discover their applicability to Industry 4.0 optimization, scheduling, risk management, and control problems. This goal reflects the present-day challenges of the dynamically growing requirements of autonomous decision support solutions, which is essential to continue to move the industrial transformation to a sustainable and environmentally friendly operation.

Reinforcement learning (RL) has a significant opportunity to revolutionize artificial intelligence (AI) applications by serving a novel approach to machine learning (ML) developments that allows the user to handle large-scale problems efficiently. These techniques, together with widely used Internet of Things (IoT) tools, have opened up new possibilities for optimizing complex systems, including domains such as logistics, project planning, scheduling, and other industry-related domains. Extracting this potential can result in fundamental progress in the transformation of Industry 4.0. Vertical and horizontal integration will be strengthened during this digital transformation, flexibility should be increased, and human control and supervision should be focused. Furthermore, the data produced by the integrated tools is growing exponentially, which requires a higher level of autonomous processes and decisions. Reinforcement learning can serve as a valuable tool in developing self-optimizing and self-organizing Industry 4.0 solutions. The main challenge of creating these applications is that several methods and techniques and a wide range of parameters need to be defined. As the definition of these parameters requires detailed knowledge of the nature of the RL algorithms, one of the main goals of my work is to provide a comprehensive overview of the RL methods from the point of view of Industry 4.0 and smart manufacturing.

The increasing environmental pressures caused by human activities drive governments and organizations to limit or reduce their footprints. As an outcome, the European Commission declared this a strategic goal. A new Circular Economy Action Plan for a cleaner and more competitive Europe has been announced to transform the economy and society sustainably. There is a wide range of recommendations and proposals on how these types of issues should be handled, and the key point is to reduce consumption and increase the recycling rate. The efforts led to the creation of the concept of circular economy, which is more complex than just manufacturing optimization because it also covers the optimization steps

of the supply chain, disassembly, and recycling. The circular economy optimization goal covers the recollecting and recycling processes that can be a bit more complex because these are more stochastic and less controlled. The most basic models assumed deterministic inputs, but the advanced models started to handle the uncertainty observed in real problems. Among others, the source materials and their distributions are described by stochastic variables. Similarly, the disassembly tasks, their required process times, and the demand values of the recycling steps can also be stochastic attributes in some models.

The supply chain optimization and assembly line balancing topics have been intensively studied issues since the 1950's years, while the focused analysis of disassembly lines started almost 40 years later. Another goal of my research was to explain in detail how to construct an effective implementation for disassembly line balancing problems.

From a broader aspect, further similar optimization problems can be found, in which formulations lead to mixed-integer (or mixed continuous-discrete) optimization problems. The next goal of my work was to identify and describe a new problem class, which is the constrained stochastic graph traversal problems. Although efficient solutions for the classical shortest pathfinding problems are known, these methods do not work with constraints and stochastically distributed distances. This motivated me to turn to the reinforcement learning method. In contrast with the most commonly used gradient-based solutions of mixed-integer optimization problems, I focused on a new discretization approach for reducing the state space into a purely discrete space representation and applied the iterative policy optimization method to that.

I identified a diverse set of problems belonging to the constrained graph traversal class. Although the shortest pathfinding problem is a well-known combinatorial optimization problem, its parameters are difficult to define precisely. It is easy to see that the optimal route problem of a truck is practically a constrained shortest pathfinding problem, where the constraints describe the working hours limits of the driver and the availability of parking or the fuel capacity limits. The daily planning of an electric delivery van is a constrained Hamiltonian pathfinding problem, where the constraints are based on the limits of the battery capacity and charging options or the opportunities for battery exchange. Furthermore, disassembling all product components after their life cycle is also a constrained graph traversal problem, where a precedence graph describes component removal dependencies and the

constraints limit the use of parallel workstations. Finally, I need to point out that, in real-life problems, these types of problems are rather stochastic optimization tasks than discrete ones.

Reinforcement learning (RL) can be an obvious solution for sequential decision-making processes such as step-by-step pathfinding. The objective function should be implemented in the reward function, as well as in violation of constraints. However, additional difficulties arise when considering stochastic effects or measurement uncertainty: some continuous components must be integrated into the state space, which becomes mixed discrete-continuous. There are successful methods to solve the mixed-integer problem using gradient-based or deep RL methods. RL techniques can also solve constrained combinatorial optimization problems. Moreover, stochastic shortest-path problems also have RL solutions. Compared to them, my research aims to discover a new algorithm by performing discretization steps using the DBSCAN clustering method and applying a discrete RL method for mixed continuous-discrete constrained stochastic graph traversal problems.

## 2 New scientific results

The following list contains my new scientific results in three thesis findings.

**1. A multidimensional classification of reinforcement learning (RL) methods for applications in Industry 4.0 facilitates the identification of prevailing research trends. Furthermore, a questionnaire-based guideline supports the adoption of an appropriate methodology for the setup of RL solutions and helps to select the most suitable RL method to address specific problems.**

1.1 I performed a systematic literature review that included keyword standardization and categorization into three major dimensions: principles captured, industrial field, and mathematical approaches to the application methodology.

1.2 I made an analysis to identify the main research trends and the correlations between the problem types, the applied RL methods, and the industrial fields.

1.3 Based on these, I created a questionnaire guideline to support the adoption of an appropriate methodology for the setup of RL solutions and help select the most suitable RL method to address specific problems.

1.4 The guideline can be improved to being a fundamental basis of an automated RL method selection application.

Related publications: [R1].

**2. The newly developed reinforcement learning-based dLOptRL algorithm can solve disassembly line balancing problems and effectively supports the optimal operation of sustainable disassembly lines.**

2.1 I prepared the general formulation of the disassembly line balancing problem for the reinforcement learning methods including the state- and action spaces and the reward function.

2.2 I developed the DLOPTRL algorithm that belongs to the Heuristically Accelerated Reinforcement Learning methods to solve disassembly line optimization problems more effectively than the original reinforcement learning method.

- 2.3 Hence there are limitations in most available linear and quadratic solvers, I reformulated the disassembly line balancing problem as a mixed-integer quadratic optimization problem using significantly fewer decision variables to provide a reference solution to evaluate the performance of the DLOPTRL algorithm.
- 2.4 I derived a step-by-step procedure for generating the RL formulation directly from the MIQP formulation.
- 2.5 I demonstrated the applicability of the DLOPTRL algorithm in small- and medium-scale disassembly line optimization use cases. I pointed out that my algorithm provides a real alternative compared to the classical Mixed-Integer Quadratic Programming solution.

Related publications: [R2], [R3].

**3. The Q-compression method is a newly developed, generally applicable reinforcement learning-based approach for solving mixed-integer optimization problems using a discrete dynamic representation of the state space across a wide range of complex optimization problems.**

- 3.1 I defined the class of restricted stochastic graph traversal problems and showed that many scheduling and routing problems belong to that. I also defined the exact method to transform the problems into a constrained stochastic shortest pathfinding problem and, hence, standardize the formulation.
- 3.2 I developed the Q-compression method to find a dynamic discrete representation of a mixed-integer optimization problem using the DBSCAN algorithm and solve it using the reinforcement learning method.
- 3.3 The Q-compression method improves the built-in heuristic of the DLOPTRL algorithm to reduce the action space and provides a human-interpretable model without time-consuming preliminary steps.
- 3.4 Finally, I demonstrated the usability of the algorithm in different selected problems and verified its advance to the grid-based discretization method.

Related publications: [R4], [R5].

### **3 Utilization of new scientific results**

The overview and questionnaire guidelines in my first thesis help researchers and developers shorten the preparation time by allowing them to focus on the most relevant methods. They can also be the basis for a later automatic method selection procedure.

My second thesis delivered the `DLOPTRL` algorithm, which offers a complete solution for disassembly line balancing problems that can learn directly from an ongoing operation or a simulated virtual twin.

My third thesis declares the constrained stochastic graph traversal problem class, which involves several essential optimization problems and can definitely be extended by further ones. This class enables the easy transfer of a solution method from one member problem to another. The Q-compression method presents a human-interpretable solution without time-consuming preliminary steps, and therefore, its implementation time can be significantly reduced. The algorithm can potentially adapt to changing environments by modifying its observation time window.

## Bibliography

Hungarian Scientific Bibliography (MTMT) personal researcher page:

<https://m2.mtmt.hu/gui2/?type=authors&mode=browse&sel=10081160>

### Related publications to theses

- [R1] Tamás Kegyes, Zoltán Süle, and János Abonyi. The applicability of reinforcement learning methods in the development of industry 4.0 applications. *Complexity*, 2021:1–31, 2021.
- [R2] Tamás Kegyes, Zoltán Süle, and János Abonyi. Disassembly line optimization with reinforcement learning. *Central European Journal of Operations Research*, pages 1–28, 2024.
- [R3] Tamás Kegyes, Zoltán Süle, and János Abonyi. Incorporation of heuristic search into q-learning. In *Short papers of VOCAL 2022 (9th VOCAL Optimization Conference: Advanced Algorithms)*, Budapest, Hungary, 2022.
- [R4] Tamás Kegyes, Alex Kummer, Zoltán Süle, and János Abonyi. Generally applicable q-table compression method and its application for constrained stochastic graph traversal optimization problems. *Information*, 15(4):193, 2024.
- [R5] Tamás Kegyes, Alex Kummer, Zoltán Süle, and János Abonyi. Improvements of the q-compression method for constrained stochastic graph traversal problems. In *10th VOCAL Optimization Conference: Advanced Algorithms: Conference Program*), Budapest, Hungary, 2024.

### Further publications

- [F1] Tamás Kegyes, Zoltán Süle, and János Abonyi. Machine learning-based decision support framework for cbrn protection. *Heliyon*, 2024.
- [F2] Kegyes Tamás, Süle Zoltán, and Abonyi János. Az információmenedzsment szerepe az abv-védelemben. *Nemzetbiztonsági Szemle*, 11(1):62–77, 2023.
- [F3] Tamás Kegyes, Zoltán Süle, and János Abonyi. Szétszerelési láncok optimalisa megerősítéses tanulással. In *Magyar Operációkutatási Konferencia: Absztraktok könyve*, Budapest, Hungary, 2021.