Reinforcement Learning Explained

Reinforcement Learning (RL) is a branch of Artificial Intelligence (AI) where agents learn optimal behavior through interaction with an environment by receiving feedback in the form of reward. After decades of research, RL has matured into a powerful technology driving real-world innovation; it is now used in areas such as robotics, energy systems, finance, and autonomous vehicles.

Yet, for many, RL feels inaccessible, buried under dense mathematics and complex theory. This book changes that. It is designed to help newcomers start applying RL as quickly as possible through a classical pedagogical approach: many small, focused examples that build intuition and practical skill step by step.

Featuring:

- Essential concepts explained from the ground up
- Code-based examples that reveal how algorithms work in practice
- Worked examples by hand to strengthen intuition, just like in engineering or mathematics textbooks
- Language-agnostic guidance, easily followed using Python, Java, or C++

Even readers without coding or university-level mathematics backgrounds will gain valuable insight into the fascinating world of RL—insight that may become a critical differentiator in the age of AI. Whether you are a student or professional, *Reinforcement Learning Explained* will give you the tools and confidence to explore one of AI's most exciting frontiers.

Reinforcement Learning Explained

A Practical Problem-Solving Approach

Jonas Hellgren and Johannes Lindgren

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Designed cover image: Cover image provided by the author.

First edition published 2026
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

Library of Congress Cataloging-in-Publication Data
[Insert LoC Data here when available]

ISBN: 9781041062264 (hbk)
ISBN: 9781032996653 (pbk)
ISBN: 9781003634409 (ebk)

DOI: 10.1201/9781003634409

Typeset in Minion
by KnowledgeWorks Global Ltd.

Contents

About the Authors

Jonas Hellgren is a researcher specializing in reinforcement learning, optimization, and electrified vehicle systems. With experience across academia and industry spanning patents, publications, thesis supervision, and industrial projects, he brings both practical insight and theoretical depth. This book reflects his commitment to making complex ideas accessible.

Johannes Lindgren is a technical consultant specializing in software development, verification, and commissioning across rail, automotive, and maritime applications. Currently at Combine, developing software for the rail sector. Previous roles include simulation and verification at Volvo Autonomous Solutions and system commissioning at Lean Marine, along with research in image segmentation at CPAC Systems.

Introduction

Reinforcement Learning (RL) is about training an **artificial agent** to **learn** from experience by trying actions, observing their consequences, and receiving rewards or penalties. Instead of following explicit instructions, the agent gradually discovers which strategies (policies) yield the best long-term outcomes. This trial-and-error process mirrors the way humans and animals learn, making RL not only powerful but also deeply intuitive.

What makes RL especially relevant today is the dramatic rise in computational power. With modern algorithms and advanced hardware, RL has evolved from a theoretical concept into a practical engine of intelligent decision-making. It now enables machines to tackle challenges that once seemed far beyond reach: robots that can walk and run with agility, vehicles that coordinate charging across smart energy grids, and algorithms that achieve—and even surpass—human expertise in some of the most complex games ever created, such as Go.

In short, RL represents a leap forward in artificial intelligence, empowering machines to adapt and innovate beyond what was imaginable just decades ago.

THE CRAFTSMEN METAPHOR

Engineers and researchers share many similarities with skilled craftsmen; both possess a variety of tools designed to yield specific results. The possession of a diverse toolkit, coupled with a deep understanding of how to effectively utilize these tools, is crucial for enhancing productivity. This book focuses on a tool valuable for engineers and researchers: RL. It is important to note that, like any tool, RL is not a one-size-fits-all solution; its true power is realized only when it is applied thoughtfully to the right challenges. As will be elaborated upon in the following chapters, RL comes in various forms and approaches. Thus, one could view RL not merely as a single tool, but rather as a versatile collection of tools, each tailored for different applications.

AMBITION

This book is partly abstract and theoretical, introducing essential concepts for understanding RL. At the same time, it is applied; by presenting and solving multiple examples, it effectively integrates theory with practice. It covers various RL techniques, from more classical methods such as Q-learning to modern approaches like actor-critic RL.

The ambition is to present the field in an engaging way without overwhelming readers with equations, theorems, and proofs. The focus is on explaining how to use the powerful RL tool rather than analyzing its inner workings. “A Beginner’s Manual for Applying Reinforcement Learning” or “Reinforcement Learning Explained by Examples” could have been alternative book titles.

WHAT IS NOT INCLUDED

Artificial intelligence refers to the simulation of human intelligence in machines programmed to think and learn. Its scope includes various applications such as robotics, natural language processing, and decision-making systems. This book covers RL, a subfield of machine learning. Machine learning, in turn, is a subfield of artificial intelligence. Other variants of machine learning include supervised learning, where the model is trained on labelled data; unsupervised learning, which involves finding patterns in unlabeled data. These other variants are not covered in detail in this book. In addition, generative AI technologies, including ChatGPT, fall outside the scope of this book.

OTHER RL TEXTBOOKS

The most well-known RL textbook is probably *Reinforcement Learning: An Introduction* [1], which covers key concepts and algorithms without focusing on a specific programming language. The authors played a crucial role in developing some of RL's foundational concepts. The book is sometimes theoretical and may be challenging for beginners.

Deep Reinforcement Learning Hands-On [2] and *Foundations of RL with Applications in Finance* [3] assume the reader has prior knowledge of Python. The focus of reference [3] is RL concepts and their applications in financial decision-making, covering topics such as portfolio optimization, algorithmic trading, and risk management. The textbook *Grokking Deep Reinforcement Learning* [4] has a focus on deep RL and includes numerous Python code examples to illustrate the concepts. *Reinforcement Learning and Optimal Control* [5] and *Control Systems and Reinforcement Learning* [6] explore RL from a control theory perspective. Compared to the books listed earlier [1]–[4], these books are more theoretical.

This book differs from earlier works in several important ways. Its later release allows it to cover state-of-the-art algorithm such as AlphaZero, which combines learning and planning. It also offers many worked examples, solved step by step "by hand", to help readers build intuition. In addition, the book includes code-based examples—some of which are designed to go deliberately wrong. By examining these cases, readers can gain valuable insights into when and how things can fail. The material has been tested on proofreaders with no prior experience in RL, with the explicit goal of making the text accessible even to complete beginners. Finally, the book is deliberately concise: it is approximately half the length of many other titles in the field, focusing only on the essentials. After all, why should a student invest in an unnecessarily thick and expensive book?

REQUIRED SKILLS AND TARGET GROUPS

To effectively read and understand this book, the reader should possess the following knowledge and skills:

Basic Mathematics: Understanding of calculus, linear algebra, and probability theory, typically introduced early in any university-level engineering education.

Programming Skills: Proficiency in at least one programming language, preferably with knowledge of object-oriented programming and clean coding practices.

Fundamental Machine Learning Knowledge: Familiarity with basic machine learning concepts and techniques, such as linear regression and classification.

Familiarity with Algorithms: Understanding of basic algorithm design and analysis. The ability to interpret pseudocode is a valuable skill.

The target audience for this book includes:

Students: Undergraduate and graduate students studying computer science, artificial intelligence, or related fields.

Researchers: Individuals conducting research in machine learning, artificial intelligence, or related disciplines.

Practitioners: Data scientists, software engineers, and professionals looking to apply RL techniques in industry.

Enthusiasts: Hobbyists and self-learners with a strong interest in artificial intelligence and machine learning.

Educators: Teachers and professors who want to incorporate RL into their curriculum.

Leaders: Industrial and academic leaders who want to gain basic knowledge of RL, perhaps to better understand the technology's potential. They can focus on the introductory chapters.

EXAMPLES AND EXERCISES

Each chapter includes **worked examples, code-based examples**, and **exercises**. The worked examples are designed to reinforce key concepts through problems solved step by step, typically "by hand". Below is a worked example (for illustrative purposes).

The code-based examples are formulated and described, but since a full solution may span thousands of lines of code, only the results are presented and explained, often with plots. In some cases, failed results are also shown to support deeper understanding. Instructions for obtaining the code are provided later in Section *External Resources*.

Exercises are placed at the end of each chapter to encourage further practice and reflection. To help readers prioritize important information, icons are sometimes used, with

WORKED EXAMPLE 1

What makes this book different from other RL books?

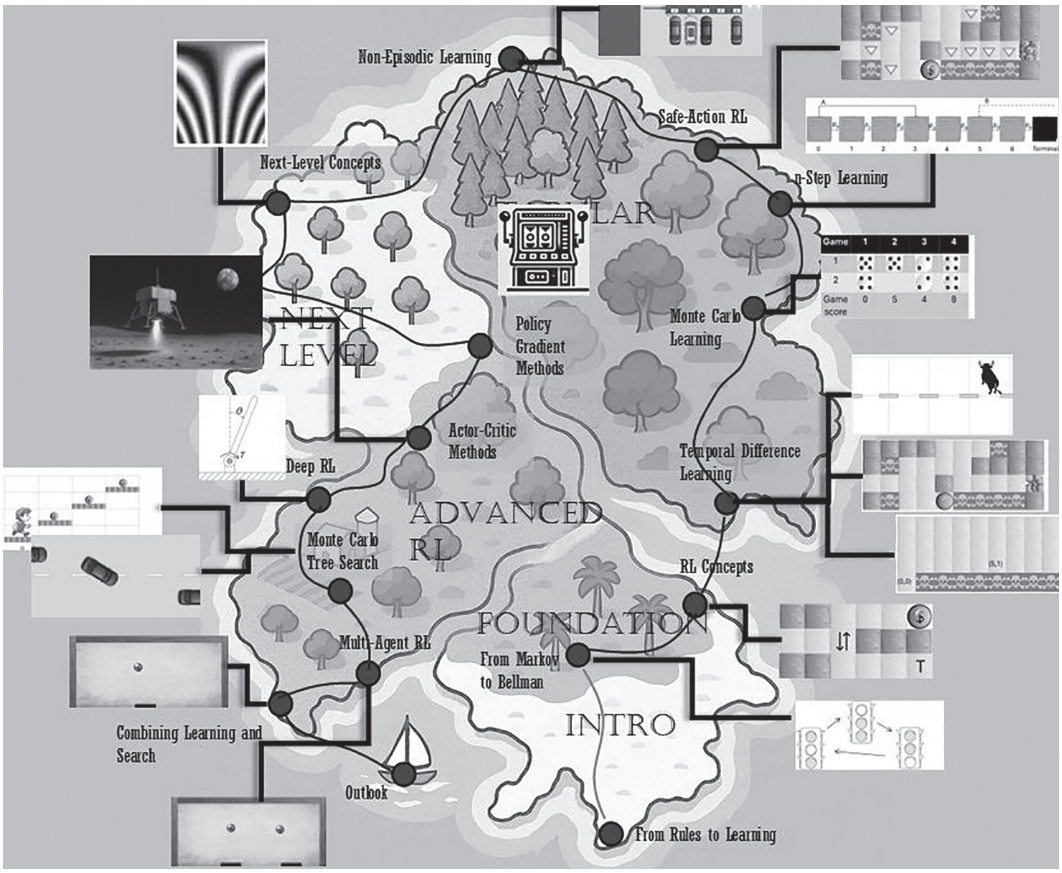
Practical—a how-to-use focus, not just theory.
Modern—includes state-of-the-art algorithms.
Accessible—tested on readers with no RL background.
Hands-on—many worked and code-based examples.
Concise—essentials covered in relatively few pages.

their meanings summarized in the table below. Throughout the book, the icons are used in all worked examples and in some equations.

Icon	Description
	Key Concept. Indicates essential material that should be prioritized. Repeat reading until you fully understand it.
	Good to Know. Provides valuable insights that enhance understanding but are not as critical as key concepts.
	Deep Dive. Covers advanced or in-depth content. It's useful but probably not strictly necessary for applying RL.
	Support. Highlights a worked example that supports understanding. It may cover familiar ground for most readers but help those who need a refresher.
	Coding Exercise. Marks an exercise that is intended to be solved through programming.

OUTLINE

The figure below shows the outline of the book. The included chapters correspond to cities in the fictive map. Related chapters can be grouped. These groups correspond to regions on the map. For example, the region *Intro* includes *From Rules to Learning* and *From Markov to Bellman*. Figures placed alongside a city (chapter) denote code-based examples.



The *Intro* region contains only the chapter *From Rules to Learning*, which motivates RL and provides historical context.

The *Foundation* region establishes the groundwork for understanding RL by introducing the Markov decision process and its essential components: states, actions, and rewards. It also presents key mathematical structures, including the Bellman equations, the value function, and discounting. Practical aspects such as parameter fitting, feature representation, and linear function approximation are covered in the chapter *From Markov to Bellman*. The chapter *Reinforcement Learning Concepts* introduces the core triad of RL: the trainer, the agent, and the environment. An illustrative example, the *Splitting Path problem*, is used to demonstrate these fundamental concepts.

The comprehensive *Tabular* region consists of five chapters. In *Temporal Difference Learning*, classical RL techniques such as Q-learning and Sarsa are presented, accompanied by worked examples including *Blocked Road Lane*, *Cliff Walk*, and *Gold Treasure*. These examples serve as practical guides for applying the methods. The *Monte Carlo Learning* chapter builds methods that use the principle of sampling returns, illustrated through the *Random Walk* and *Dice Game* problems. The chapters *n-Step Learning* and *Safe-Action Reinforcement Learning* both focus on techniques for accelerating learning. Finally, *Non-Episodic Learning* introduces an alternative formulation in which the training process is continuous and never terminates.

In the region *Next Level*, there is only one chapter: *Next-Level Concepts*. It prepares for more advanced RL methods. The dimensionality of real-world problems is often demanding, with numerous input variables and an immense number of possible states. To address this complexity, the chapter introduces core ideas that underpin more advanced RL methods. A central theme is gradient descent learning, a fundamental tool in machine learning for training function approximators. In addition, the chapter explores two influential function approximators: radial basis function networks and neural networks. Through clear, illustrative examples, readers are guided in applying these techniques effectively.

The following chapters are included in the region *Advanced RL*: *Policy Gradient Methods*, *Actor-Critic Methods*, *Deep Reinforcement Learning*, *Monte Carlo Tree Search*, *Combining Learning and Search,* and *Multi-Agent Reinforcement Learning*. Compared to the chapters in the *Tabular* region, these chapters introduce techniques that promote more stable training and handle problems with richer or continuous state spaces. One illustrative problem in Policy Gradient Methods is the *Cannon Shooting* problem, where the challenge is to set the firing angle (a continuous variable) to hit a target. Other presented and solved problems include landing a lunar lander safely on the moon's surface and balancing an inverted pendulum.

The book closes with an *Outlook* chapter, reflecting on the future of RL and how the field is expected to evolve. As suggested by the map on the previous page, this chapter does not rest on solid ground but rather ventures out onto the open ocean, where much remains uncharted.