Deep Reinforcement Learning

Mohit Sewak

Deep Reinforcement Learning

Frontiers of Artificial Intelligence

Springer

Mohit Sewak
Pune, Maharashtra, India

ISBN 978-981-13-8284-0 ISBN 978-981-13-8285-7 (eBook)
https://doi.org/10.1007/978-981-13-8285-7

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Reinforcement Learning has evolved a long way with the enhancements from deep learning. Recent research efforts into combining deep learning with Reinforcement Learning have led to the development of some very powerful *deep Reinforcement Learning* systems, algorithms, and agents which have already achieved some extraordinary accomplishment. Not only have such systems surpassed the capabilities of most of the classical and non-deep-learning-based Reinforcement Learning agents, but have also started outperforming the best of human intelligence at tasks which were believed to require extreme human intelligence, creativity, and planning skills. Some of the DQN-based agents consistently beating the best of human players at the complex game of AlphaGo are very good examples of this.

This book starts with the basics of Reinforcement Learning and explains each concept using very intuitive and easy to understand examples and applications. Continuing with similar examples, this book then builds upon to introduce some cutting-edge researches and advancements that make Reinforcement Learning outperform many of the other (artificial) intelligent systems. This book aims to not only equip the readers with just the mathematical understanding of multiple cutting-edge Reinforcement Learning algorithms, but also prepares them to implement these and similar advanced *Deep Reinforcement Learning* agents and system hands-on in their own domain and application area.

This book starts from the basic building blocks of Reinforcement Learning, then covers the popular classical DP and classical RL approaches like value and policy iteration, and then covers some popular traditional Reinforcement Learning algorithms like the TD learning, SARSA, and the Q-Learning. After building this foundation, this book introduces deep learning and implementation aids for modern Reinforcement Learning environments and agents. After this, the book starts diving deeper into the concepts of *Deep Reinforcement Learning* and covers algorithms like the deep Q networks, double DQN, dueling DQN, (deep) synchronous actor-critic, (deep) asynchronous advantage actor-critic, and the deep deterministic policy gradient. Each of the theoretical/mathematical chapters on these concepts is followed by a chapter on practical coding and implementation of these agents' grounds-up connecting the concepts to the code.

Pune, India Mohit Sewak

Who This Book Is For?

This book will equally appeal to readers with prior experience in deep learning, who want to learn new skills in Reinforcement Learning, and to readers who have been the practitioner of Reinforcement Learning or other automation systems and who want to scale their knowledge and skills to *Deep Reinforcement Learning*. By combining the concepts from deep learning and Reinforcement Learning, we could come closer to realizing the true potential of ‘general artificial intelligence’.

Besides presenting the mathematical concepts and contemporary research in the field of *Deep Reinforcement Learning*, this book also covers algorithms, codes, and practical implementation aids for both Reinforcement Learning environments and agents. This book is intended to be a guide and an aid for both the types of readers, for the ones who are interested in the academic understanding and being abreast with some of the latest advancements in *Deep Reinforcement Learning* and also for the ones who want to implement these advanced agents and systems into their own fields.

Ranging from application in autonomous vehicles to dynamic scheduling and management of production process, to intelligent maintenance of critical machineries, to driving efficiency in utility management, to making automated systems for health care, to intelligent financial trading and transaction monitoring, to aiding intelligent customer engagement, and to mitigating high-throughput cyber threats, the concepts learnt in this book could be applied to multiple fields of interest.

The code in the book is in Python 3x. The deep learning part of the code uses the TensorFlow library. Some code also uses the Keras wrapper to TensorFlow. *Deep Reinforcement Learning* wrappers like Keras-RL are also demonstrated. This book expects basic familiarization in Python with object-oriented programming concepts to enable implementation of distributed and scalable systems.

What This Book Covers?

Chapter 1—*Introduction to Reinforcement Learning*—covers the basic design of Reinforcement Learning and explains in detail the concepts like the environment, actor, state, and rewards, and the challenges in each.

Chapter 2—*Mathematical and Algorithmic Understanding of Reinforcement Learning*—builds upon a strong mathematical and algorithmic foundation to understand the internal functioning in different types of agents.

Chapter 3—*Coding the Environment and MDP Solution*—illustrates how to build a custom Reinforcement Learning environment in code over which different reinforcement agents can train and also implements the value iteration and policy iteration algorithms over a custom environment.

Chapter 4—*Temporal Difference Learning, SARSA, and Q-Learning*—covers the TD learning estimation process and the on-policy SARSA and off-policy Q-Learning algorithms along with different types of exploration mechanism.

Chapter 5—*Q-Learning in Code*—implements the Q-Learning algorithm in Python via the tabular approach using the epsilon-greedy algorithm for behavior policy.

Chapter 6—*Introduction to Deep Learning*—introduces the concepts of deep learning like layer architecture, activation, loss functions, and optimizers for the MLP-DNN and CNN algorithms.

Chapter 7—*Implementation Resources*—covers the different types of resources available to implement, test, and compare cutting-edge deep Reinforcement Learning models and environments.

Chapter 8—*Deep Q Network (DQN), Double DQN, and Dueling DQN*—covers the deep Q networks and its variants the double DQN and the dueling DQN and how these models surpassed the best of human adversaries' performance at the game of AlphaGo.

Chapter 9—*Double DQN in Code*—covers implementation of a double DQN with an online active Q network coupled with another offline target Q network with both networks having customizable deep learning architecture, built using Keras on TensorFlow.

Chapter 10—*Policy-Based Reinforcement Learning Approaches*—covers the basic understanding of policy-based Reinforcement Learning approaches and explains the policy-gradient mechanism with the reinforce algorithm.

Chapter 11—*Actor-Critic Models and the A3C*—covers stochastic policy-gradient-based actor-critic algorithm with its different variants like the one using 'advantage' as a baseline and those that could be implemented in 'synchronous' and 'asynchronous' distributed parallel architectures.

Chapter 12—*A3C in Code*—covers the implementation of the asynchronous variant of the distributed parallel actor-critic mechanism with multiple agents working simultaneously to update the master's gradient. The agent algorithm is implemented using the TensorFlow library, using the libraries' 'eager execution' and model's 'sub-classing' features.

Chapter 13—*Deterministic Policy Gradient and the DDPG*—covers the deterministic policy-gradient theorem and the algorithm and also explains the enhancements made to enable the deep learning variant of deep deterministic policy gradient (DDPG).

Chapter 14—*DDPG in Code*—covers the implementation of the DDPG algorithm to enable the Reinforcement Learning tasks requiring continuous-action control and implements it in a very few lines of code using the Keras-RL wrapper library.

Contents

About the Author

Mohit Sewak is a Ph.D. scholar in CS&IS (Artificial Intelligence and Cyber Security) with BITS Pilani - Goa, India, and is also a lecturer on subjects like Artificial Intelligence, Machine Learning, Deep Learning and NLP for the post-graduate technical degree program. He holds several patents (USPTO & Worldwide) and publications in the field of Artificial Intelligence and Machine Learning.

Besides his academic linkages, Mohit is also actively engaged with the industry and has many accomplishments while leading the research and development initiatives of many international AI products. Mohit has been leading the Reinforcement Learning practice at QiO Technologies, the youngest player in Gartner's magic quadrant for Industry 4.0.

In his previous roles, Mohit had led the IBM Watson Commerce in India's innovation initiative in cognitive line of feature as Sr. Cognitive Data Scientist. Mohit had also been the Principal Data Scientist for IBM's global IoT products like IBM Predictive Maintenance & Quality. He had also been the advanced analytics architect for IBM's SPSS suite in India.

Mohit has over 14 years of very rich experience in researching, architecting and solutioning with technologies like TensorFlow, Torch, Caffe, Theano, Keras, Open AI, OpenCV, SpaCy, Gensim, Spark, Kafka, ES, Kubernetes, and Tinkerpop.