Learning to Play

Aske Plaat

Learning to Play

Reinforcement Learning and Games

Springer

Aske Plaat
Leiden Institute of Advanced Computer Science
Leiden University
Leiden, The Netherlands

ISBN 978-3-030-59237-0 ISBN 978-3-030-59238-7 (eBook)
https://doi.org/10.1007/978-3-030-59238-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my students

Contents

Preface

Amazing breakthroughs in reinforcement learning have taken place. Computers teach themselves to play Chess and Go and beat world champions. There is talk about expanding application areas towards general artificial intelligence (AI). The breakthroughs in Backgammon, Checkers, Chess, Atari, Go, Poker, and StarCraft have shown that we can build machines that exhibit intelligent game playing of the highest level. These successes have been widely publicized in the media, and inspire AI entrepreneurs and scientists alike. Reinforcement learning in games has become a mainstream AI research topic. It is a broad topic, and the successes build on a range of diverse techniques, from exact planning algorithms, to adaptive sampling, deep function approximation, and ingenious self-play methods.

Perhaps because of the breadth of these technologies, or because of the recency of the breakthroughs, there are few books that explain these methods in depth. This book covers all methods in one comprehensive volume, explaining the latest research, bringing you as close as possible to working implementations, with many references to the original research papers.

The programming examples in this book are in Python, the language in which most current reinforcement learning research is conducted. We help you to get started with machine learning frameworks such as Gym, TensorFlow, and Keras, and provide exercises to help understand how AI is learning to play.

This is not a typical reinforcement learning textbook. Most books on reinforcement learning take the single-agent perspective, of path finding and robot planning. We take as inspiration the breakthroughs in game playing, and use two-agent games to explain the full power of deep reinforcement learning.

Board games have always been associated with reasoning and intelligence. Our games perspective allows us to make connections with artificial intelligence and general intelligence, giving a philosophical flavor to an otherwise technical field.

Artificial Intelligence

Ever since my early days as a student I have been captivated by artificial intelligence, by machines that behave in seemingly intelligent ways. Initially I had been taught that, because computers were deterministic machines, they could never do something new. Yet in AI these machines do complicated things such as recognize patterns, and play Chess games. Actions emerged from these machines, behavior that appeared not to have been programmed into them. The actions seemed new, and even creative, at times.

For my thesis I got to work on game playing programs for combinatorial games such as Chess, Checkers, and Othello. The paradox became even more apparent. These game playing programs all followed an elegant architecture, consisting of a search function and an evaluation function.[1] These two functions together could find good moves all by themselves. Could intelligence be so simple?

The search-evaluation architecture has been around since the earliest days of computer Chess. Together with minimax, it was proposed in a 1952 paper by Alan Turing, mathematician, code-breaking war hero, and one of the fathers of computer science and artificial intelligence. The search-evaluation architecture is also used in Deep Blue, the Chess program that beat World Champion Garry Kasparov in 1997 in New York.

After that historic moment, the attention of the AI community shifted to a new game with which to further develop ideas for intelligent game play. It was the East Asian game of Go that emerged as the new grand test of intelligence. Simple, elegant, and mind-bogglingly complex.

This new game spawned the creation of important new algorithms, and not one, but two, paradigm shifts. The first algorithm to upset the worldview of games researchers was Monte Carlo Tree Search, in 2006. Since the 1950s generations of game playing researchers, myself included, were brought up with minimax. The essence of minimax is to look ahead as far as you can, to then choose the best move, and to make sure that all moves are tried (since behind seemingly harmless moves deep attacks may hide that you can only uncover if you search all moves). And now Monte Carlo Tree Search introduced randomness into the search, and sampling, deliberately missing moves. Yet it worked in Go, and much better than minimax.

Monte Carlo Tree Search caused a strong increase in playing strength, although not yet enough to beat world champions. For that, we had to wait another ten years.

In 2013 our worldview was in for a new shock, because again a new paradigm shook up the conventional wisdom. Neural networks were widely viewed to be too slow and too inaccurate to be useful in games. Many Master's theses of stubborn students had sadly confirmed this to be the case. Yet in 2013 GPU power allowed the use of a simple neural net to learn to play Atari video games just from looking at the video pixels, using a method called deep reinforcement learning. Two years and much hard work later, deep reinforcement learning was combined with Monte Carlo

[1] The search function simulates the kind of look-ahead that many human game players do in their head, and the evaluation function assigns a numeric score to a board position indicating how good the position is.

Tree Search in the program AlphaGo. The level of play was improved so much that a year later finally world champions in Go were beaten, many years before experts had expected that this would happen. And in other games, such as StarCraft and Poker, self-play reinforcement learning also caused breakthroughs.

The AlphaGo wins were widely publicized. They have had a large impact, on science, on the public perception of AI, and on society. AI researchers everywhere were invited to give lectures. Audiences wanted to know what had happened, whether computers finally had become intelligent, what more could be expected from AI, and what all this would mean for the future of the human race. Many start-ups were created, and existing technology companies started researching what AI could do for them.

The modern history of computer games spans some 70 years. There has been much excitement. Many ideas were tried, some with success. Games research in reinforcement learning has witnessed multiple paradigm shifts, going from heuristic planning, to adaptive sampling, to deep learning, to self-play. The achievements are large, and so is the range of techniques that are used. We are now at a point where the techniques have matured somewhat, and achievements can be documented and put into perspective.

In explaining the technologies, I will tell the story of how one kind of intelligence works, the intelligence needed to play two-person games of tactics and strategy. (As to knowing the future of the human race, surely more is needed than an understanding of heuristics, deep reinforcement learning, and game playing programs.) It will be a story involving many scientists, programmers, and game enthusiasts, all fascinated by the same goal: creating artificial intelligence. Come and join this fascinating ride.

Acknowledgments

This book would not have been possible without the help of many friends. First of all, I want to thank everybody at LIACS, for creating such a wonderful and vibrant CS & AI department to work in. So many people make this place such a joy to do research; I love you all!

Many people have read and commented on drafts of this book. Thank you Vedran Dunjko, Mike Huisman, Walter Kosters, Wojtek Kowalczyk, Alfons Laarman, Thomas Moerland, and Mike Preuß for your very sharp eyes and criticism. Thank you to Andrius Bernatavicius, Hui Wang, Zhao Yang, Michael Emmerich, Wojtek Kowalczyk, Mike Preuß, Thomas Moerland, Xuhan Lu, Joost Broekens, Matthias Müller-Brockhausen, my friends and fellow conspirators in the reinforcement learning group. In addition, thank you to Lise Stork and Wojtek Kowalczyk for discussions on deep learning and zero-shot learning. Thanks to Michael Emmerich for curriculum learning. Thank you to Walter Kosters, Frank Takes, Jan van Rijn, and Jonathan Vis for discussions on teaching, MCTS, and combinatorial games. Thanks to Mike Preuß for countless discussions and for running the games research at Leiden. Thanks to Max van Duijn for inspiring discussions on theory of mind and creativity. Thank you Walter Kosters for a shared love of combinatorics and a sense of humor. A very special thank you to Thomas Bäck, with whom I have had so many discussions on science, the universe, and everything. Without you, this book would not have been possible.

I thank all students of the reinforcement learning course in Leiden, past, present, and future, for their wonderful enthusiasm, sharp questions, and many suggestions.

Finally, I thank Saskia, Isabel, Rosalin, and little Lily, for being who they are, for knowing no limits, for playing games, for enduring me, and for their boundless love.

Leiden,
April 2020 *Aske Plaat*