Challenge for the Reader

We strongly encourage you to try training tic-tac-toe models for yourself! Note that this example is more involved than other examples in the book, and will require greater computational power. We recommend a machine with at least a few CPU cores. This requirement isn't too onerous; a good laptop should suffice. Try using a tool like `htop` to check that the code is indeed multithreaded. See how good a model you can train! You should be able to beat the random baseline most of the time, but this basic implementation won't give you a model that always wins. We recommend exploring the RL literature and expanding upon the base implementation to see how well you can do.

Review

In this chapter, we introduced you to the core concepts of reinforcement learning (RL). We walked you through some recent successes of RL methods on ATARI, upside-down helicopter flight, and computer Go. We then taught you about the mathematical framework of Markov decision processes. We brought it together with a detailed case study walking you through the construction of a tic-tac-toe agent. This algorithm uses a sophisticated training method, A3C, that makes use of multiple CPU cores to speed up training. In Chapter 9, you'll learn more about training models with multiple GPUs.