

Paper summary: Rainbow: combining improvements in deep reinforcement learning

Marko Guberina 19970821-1256

March 26, 2022

- 1 Idea in few sentences
- 2 Explanation of the central concept
- 3 Methodology
- 4 Initial rambly notes

It's a DeepMind paper in which they combine every improvement made to DQN and see what specific combination is the best. The graph showing the performance immediately shows you why you want this: it's DQN, but a LOT faster and practically 2 times better than any improvement in isolation.

4.1 Abstract

The paper combines all improvements made to DQN and empirically finds a good combination of all improvements.

4.2 Introduction

You have DQN. It has problems. People introduced the following improvements:

1. **double DQN (DDQN)**: address an overestimation bias of Q-learning by decoupling selection and evaluation of the bootstrap action.
2. **prioritized experience replay**: improves data efficiency by replaying the transitions from which there is more to learn more often
3. **dueling network architecture** helps generalize across actions by separately representing state values and action advantages
4. learning from multi-step bootstrap targets like in A3C shifts the bias-variance trade-off and helps to propagate newly observed rewards to earlier states faster

5. **distributional Q-learning** learns a categorical distribution of discounted returns instead of estimating the mean
6. **noisy DQN** uses stochastic network layers for exploration
7. many more ...

These various approaches address different issues in DQN. Some have been combined, like prioritized DDQN and dueling DDQN (which can also have prioritized experience replay). This paper combines everything.

4.3 Method

4.4 Other stuff