Review of Playing Atari with Deep Reinforcement Learning

General outline

Google DeepMind's team is set out to play 7 Atari game classics via a convolutional neural network. The agent takes the same inputs as a human player (pixels on a screen). Challenges faced by the team were finding the correct heuristics (large amounts of unlabelled data, long timeframes between action and reward) and shifting data structures as the agent learns.

The full paper can be accessed here.

The approach

The team uses a sequence of observations, actions and rewards to have the agent interact with its environment. At each time-step selects an action from a set of legal actions based on the expected reward and current state. As actions and rewards can be spaced out over long periods, the agent stores long sequences of actions and uses Markov decision processes to evaluate the best action.

At each timestep, the agent's "experience" so far is stored and fed into a "replay memory". This replay memory in return creates sequences that are then drawn from in future play-throughs. Via Deep Reinforcement learning (deep Q-learning), these experiences and chosen sequences are then refined over time.

To improve the standard Q-learning approach, the team allows each step of experience to be used in many weighted updates, to give greater data efficiency. It also breaks up correlations between experience steps to not learn from consecutive samples and increase efficiency. Also, they took additional steps to not fall into divergence traps by smoothing out the learning experience.

In order to make the input computationally easier, the initial 210x160 pixel image with 128 colours is reduced to an 84x84pixel playing field in greyscale. Using this grey-scaling turned out to be a hindrance in one special example, where due to the loss of colour, a harmful object in the game world was fully hidden when trying to reproduce results.

The results

Using this approach on 7 different Atari games without changing the architecture, the results are promising. Of course, available actions and reward definitions had to be individually altered. On all 7 games, the deep learning model outperformed previous AI agents and it was better than human experts on 3 of them.

The resulting DQN (Deep Q-network) approach gives a fresh and usable model for reinforcement learning problems.

Results have been published and recreated in many different environments.