Summary of the paper: Playing Atari with Deep Reinforcement Learning

Goals and techniques

The paper introduces the first model which connects a reinforcement learning algorithm to a deep neural network. This model was able to play successfully Atari games.
The main idea was to let the model learn how to play by capturing the Atari screen such as a human player would learn to play.

The authors describe two main goals on the paper. The first one was to create an agent able to play as many games as possible. That is, the agent should be able to generalize as many problems as possible and to solve them. The second one was to join both reinforcement learning and deep neural network together.

The model uses a Convolutional Neural Network (CNN) which is trained under a variant of a Q-learning, a stochastic gradient descent updates and a technique called *experience replay*.

The Atari frames were preproced before inputting it on the CNN in order to reduce the input dimensionality. The steps were converting the RGB to a gray-scale, down-sampling the input image from 210 x 160 to 110 x 84 pixels and finally cropping an 84 x 84 region.

The CNN, which extracts the features directly from the Atari raw pixels, has two hidden layers and two fully-connected layers. The output layer (which is one of the fully-connected layers) maps to a single action on a set of valid actions.
After each action, the transition is stored in a *experience replay* dataset which will be used a RMSProp algorithm with minibatches. After performing the *experience replay*, the model selects an action according to an ε-greedy policy.

Although the experience replay is limited by the memory buffer and the recent transitions, the model was able to overcome since the behavior distribution is averaged over many of its previous states.

Results

There were used 7 games on the experiments. In order to test the agent's generalization ability, the same architecture, learning algorithm and hyperparameters settings were applied on all of them.
The model was able to outperform previous approaches on all of the 7 games and surpass human expert on three of them.
However, there were games that the human performance did better. That is because the games were more challenging and it was necessary to the neural network to find a strategy that extends over long time.