

Summary of “Playing Atari with Deep Reinforcement Learning”

Alex Staravoi tau

alex.staravoi tau@gmail.com

Introduction

The “Playing Atari with Deep Reinforcement Learning” paper [1] describes a slightly different approach on game playing, as it uses reinforcement learning based solely on visual input, having no access to internal game state whatsoever. Not only this approach simulates the way a human player would learn to play a game, it has proven itself to be transferable to a number of completely different games on Atari platform without any fine-tuning or game-specific adjustments.

Approach

The approach described in the paper uses reinforcement learning in order to train a model to play a number of games on Atari platform. They claim not to modify any of the model hyper-parameters and not to apply any game specific optimisations, thus proving that the approach they have taken is generalisable and transferable to a variety of tasks.

Model and training

The model described in this paper is a convolutional neural network, whose input is raw video data (e.g. pixel values of the Atari screen), and whose output is a value function estimating future rewards. It can take a set of allowed game actions as controls, and the reward a model is trying to maximise is the actual game score, so the training process is highly correlated with the way a human would approach learning to play a new game. The way learning process goes is that the model takes random actions at first accumulating a so-called *replay memory* which stores last ‘N’ experiences [1, p. 4], where experience is a tuple of previous state, the action taken, reward received and the new state. As the model accumulates experiences it also per-

forms gradient descent steps on the estimator weights, based on the reward it received in a random sample from past experiences [1, p. 5]. Over time the model is tuned to be less likely to fallback to a random action and take a meaningful step instead [1, p. 7], where meaningful means maximising potential gain in the long run.

Challenges

Convolutional networks have proven themselves to be powerful feature extractors, however, there are some challenges associated with using convolutional network based models for reinforcement learning.

- **Correlated data samples during training.** In most cases deep learning models don't expect any correlation between inputs and assume there is no surrounding context, e.g. each sample is independent. Reinforcement learning, however, typically works with sequences of highly correlated data. The proposed *replay memory* approach mentioned earlier is there to address exactly this issue: as it randomly samples previous transitions, it makes the distribution of training samples more smooth over past experiences [1, p. 1].
- **Delay between actions and rewards.** Deep learning models usually represent mapping of input features to output values, whereas in environment described in this paper (e.g. Atari game) an agent may not receive a reward until many steps in the future. Basically the task paper authors are trying to solve in this paper is only partially observed, and it is impossible to come up with a robust action policy by looking at individual samples. In order to address that they consider sequences of actions and observations, rather than independent examples. This allows the authors to use standard reinforcement learning techniques by considering a sequence of actions and observations as a single state representation at some point in time [1, p. 1].

Results

Authors have proven that their model can learn to play a variety of games without changing its architecture or any hyper-parameters. Not only it reached state-of-the-art results for 6 out of 7 games it was evaluated on, it outperforms a human expert on 3 of those games. As per limitations of their model, it fails to capture long-term

strategies due to its capacity, hence can not reach human expert performance in games where such strategies are important [1, p. 8].

References

1. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (Dec 2013). *Playing Atari with deep reinforcement learning*. Technical Report arXiv:1312.5602 [cs.LG], Deepmind Technologies.