NEURAL EPISODIC CONTROL

This paper is written by a team of DeepMind, a Google company based in London and published on arXiv repository[1] on March 6, 2017. This paper attempts to find a better deep reinforcement method for AI game-playing agents.

Traditionally, reinforcement learning and deep learning prove to be successful at learning and at times performing better than human. However, the process of learning a wide-range of game environments is highly inefficient and requires enormous amount of computing power. One example is game-playing on Atari 2600 in which computers require more 200 hours in order to achieve similar results for human to achieve in 2 hours.

The slow and inefficient learning deep Q-learning process is due to several factors, such as: (i) stochastic gradient descent requires the use of small learning rates, (ii) big game environments with sparse rewards can be difficult to learn and makes prediction highly inaccurate, and finally (iii) back-propagation process in uncorrelated mini-batches can be painstakingly slow.

To improve the performance of deep Q-learning, the DeepMind team proposes to find a better solution to address the three challenges above. Inspired by the role of Hippocampus in human brain, the team suggests to use semi-tabular representation to remember the network's experience of learning the new environment. This new method is named Neural Episodic Control (NEC).

Unlike LSTM model, which reads, writes, and disposes memory, as the network learns, NEC model writes all experiences to the memory and allows it to grow very large compared to existing memory architecture. Reading from this network is made efficient using kd-tree based nearest neighborhoods. The architecture of NEC model adds a memory module known as Differentiable Neural Dictionary (DND). This DND unit has two operants module, which is "lookup" and "write". The lookup value of DND unit is a weighted sum of the values in the memory. To make DND unit scalable, the DeepMind team uses the top p-nearest neighborhoods (typically p=50) and an approximate kd-tree nearest neighbors algorithms to perform the lookup.

The result of NEC model outperforms any other algorithms to play 57 popular Atari games, such as Bowling, Pac-Man, Alien, Boxing and Pong. Across most games, NEC is significantly faster at learning in the initial phase and achieve human-level performance in about 25% of the games within 10 million frames. The learning curve on Alien game, NEC scores at 4,000 points compared to 1,800 points for the traditional Deep Q-learning algorithm and less than 1,000 points for other algorithms.

---

[1] Pritzel, Uria, Srinivasan, Puigdomenech, Vinyals, Wierstra, Blundell, NEURAL EPISODIC CONTROL, arXiv:1703.01988v1, March 6, 2017.