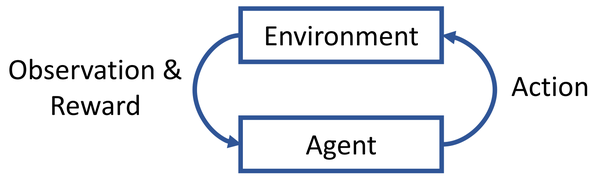
Transfer Learning On A Reinforcement Learning Application

Vishal V Kole - vvk3025@rit.edu
Advisor: Dr. Zack Butler - zxbvcs@rit.edu
Department of Computer Science | Rochester Institute of Technology

Goal

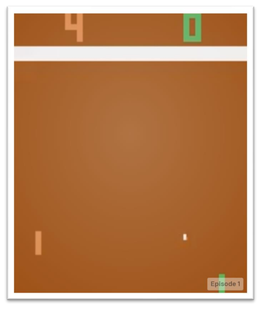
This work focuses to study transfer learning on a set of applications, the solutions to which perform better when reinforcement learning is used.

Reinforcement Learning

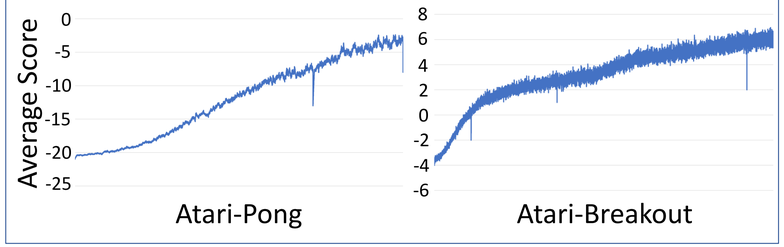


- Agent reacts with the environment and performs an action. The environment returns observation (environment state) and rewards

Environments and Agents

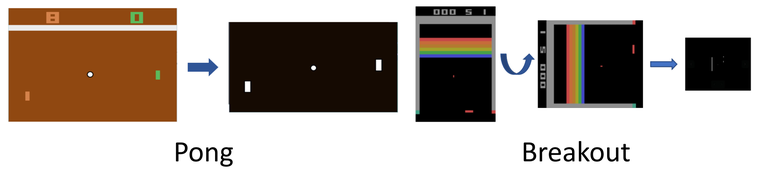


- Atari-Pong (top-left), was trained on 40,000 games, scoring 18 points on average, while Atari-Breakout (top-right) was trained on 5,80,000 games, scoring 11 points on average



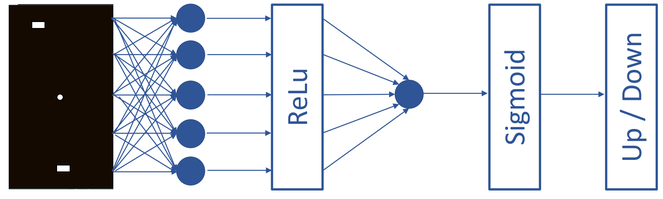
Architecture

➢ Preprocessing
- Game state is an image of the game from the simulator
- The input is cropped, down sampled, and filtered



➢ Use Of History
- Agent makes use of the previous position of the ball
- frames are merged to create a new input frame

➢ Neural Network
- A simple two layer architecture with 200 neurons in the first layer
- Activation function after the first layer is ReLU, and after the final layer is Sigmoid

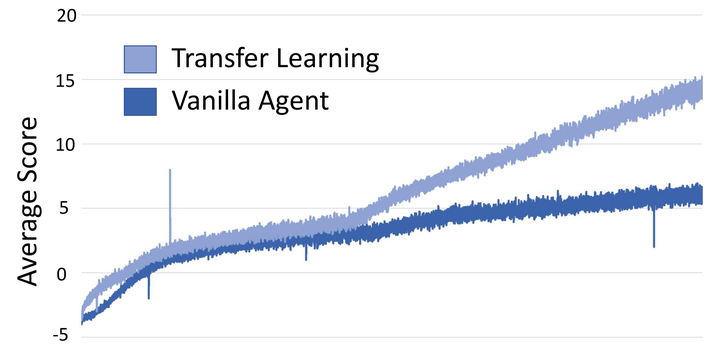


➢ Randomness And Delayed Rewards
- Some of the steps in the environment are random
- You encounter all the rewards at the end of the game, this is known as delayed reward problem

➢ Updating the Weights
- We use gradient and loss from the episodes and use RMSProp, a gradient descent algorithm to train
- Penalizes the agent if it loses and promotes if it wins

Transfer Learning Agent Analysis

- To transfer the knowledge from Pong to Breakout, we used the weights we had after training the Pong agent
- The graph shows a high learning rate at the beginning of the learning phase, this is because of the moves it learned on the game of Pong.
- After the game number 2,70,000 the agent started to get a distinctively higher average score and it kept increasing linearly.



Transfer Learning comparison on Atari-Breakout

Conclusions

- We were successfully able to transfer the knowledge from one agent to another to make it learn better.
- We also found that doing so in the domain of reinforcement learning does give us an advantage.
- Similar scoring function and dynamics helps to transfer the knowledge effectively

References

[1] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M. Riedmiller. Playing Atari with Deep Reinforcement Learning CoRR, vol. abs/1312.5602, 2013.
[2] Bellemare, M. G., et al. The Arcade Learning Environment: An Evaluation Platform for General Agents. Journal of Artificial Intelligence Research, vol. 47, 2013, pp. 253279.