FIT5190 Introduction to IT Research Methods

Deep reinforcement learning and its applications to video games

Southeast University – Monash University Joint Graduate School

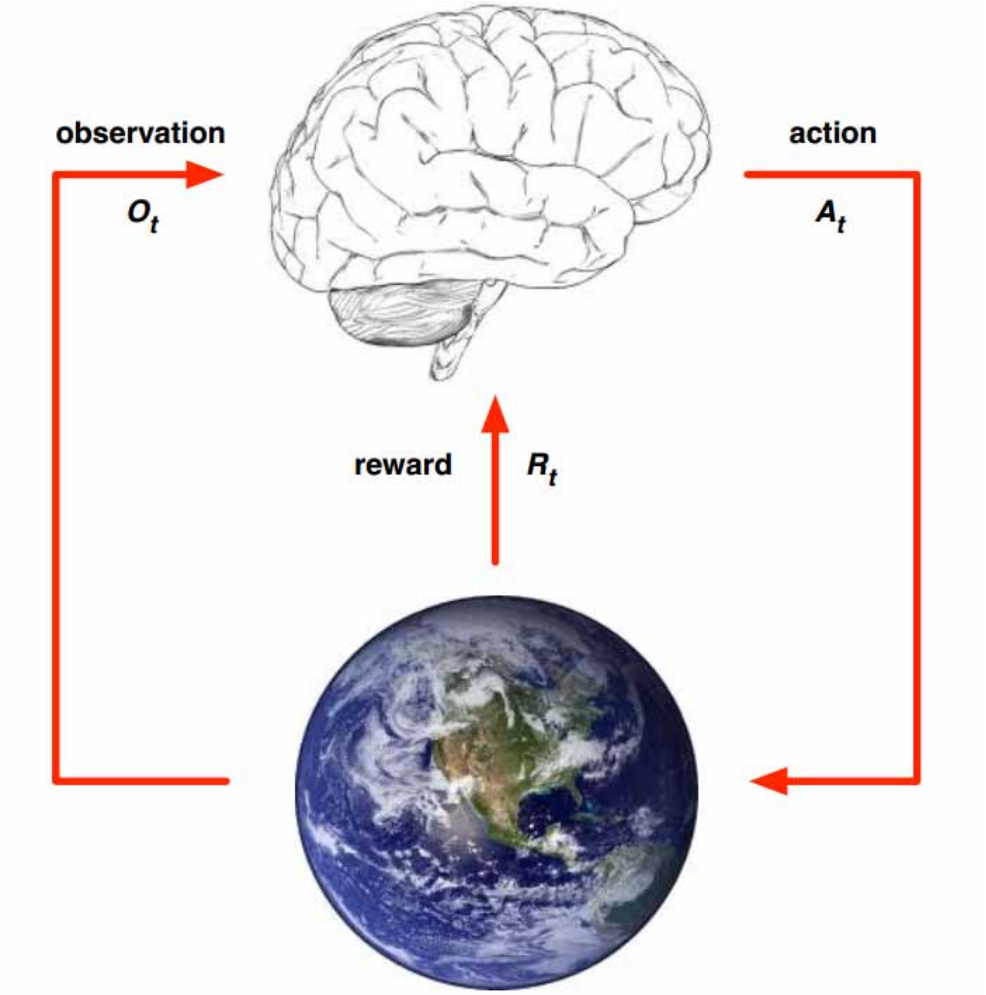
Li Shangyi
2819****

CONTENTS

1 Deep Reinforcement learning (DRL)

Insights of reinforcement learning
What is deep reinforcement learning

Reinforcement learning



- At each time step the agent:
  - Executes action A_t
  - Receives observation O_t
  - Receives scalar reward R_t
- The environment:
  - Receives action A_{t+1}
  - Emits observation O_{t+1}
  - Emit scalar reward R_{t+1}

Reinforcement learning

↓↘→↓↘→+A

Reward in reinforcement learning

- Reward:
  - The game of go
    - + reward for winning game
    - - reward for losing a game
    - 0 reward during a game
  - Robot walk
    - + reward for forward motion
    - - reward for falling over
  - Stock trading
    - + reward for making money
    - - reward for losing money

Reinforcement learning

- Sequential decision making

  $S_t, R_t, A_t, S_{t+1}, R_{t+1}, A_{t+1}, \ldots$

- Goal: select actions to maximize total future reward (cumulative reward)

Components of an agent

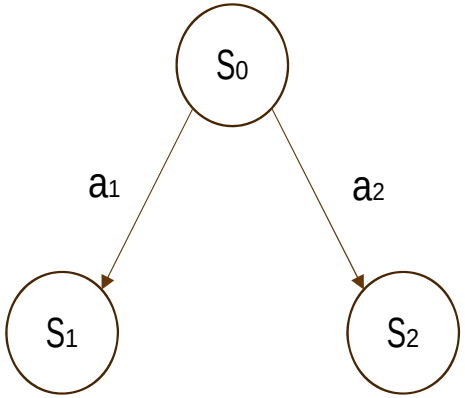
- Policy
  - Agent's behavior function
  - A map from state to action
- Value function
  - how good is each state and / or action

$$v_\pi(s) = \mathbb{E}_\pi \left[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots \mid S_t = s \right]$$

Starting from state s, then following policy π

$$q_\pi(s, a) = \mathbb{E}_\pi \left[R_{t+1} + \gamma q_\pi(S_{t+1}, A_{t+1}) \mid S_t = s, A_t = a \right]$$

Starting from state s, taking action a, then following policy π



Deep Reinforcement learning

- deep learning + reinforcement learning
  - using deep neural networks to approximate value function or policy
- No explicit training data

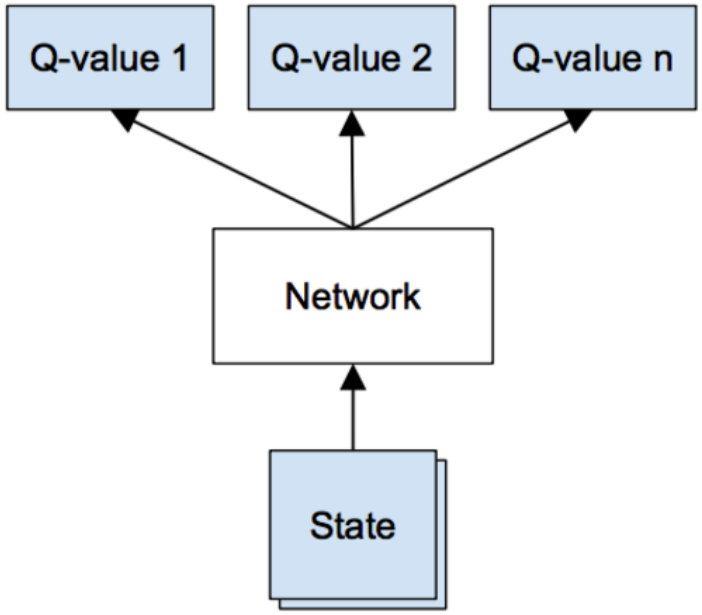
2 Applications in video games

Atari Games (Mnih et al., 2015)

https://www.youtube.com/watch?v=V1eYniJ0Rnk

https://www.youtube.com/watch?v=rz2qWeMaqtw

Atari Games



Deep Q network (DQN)

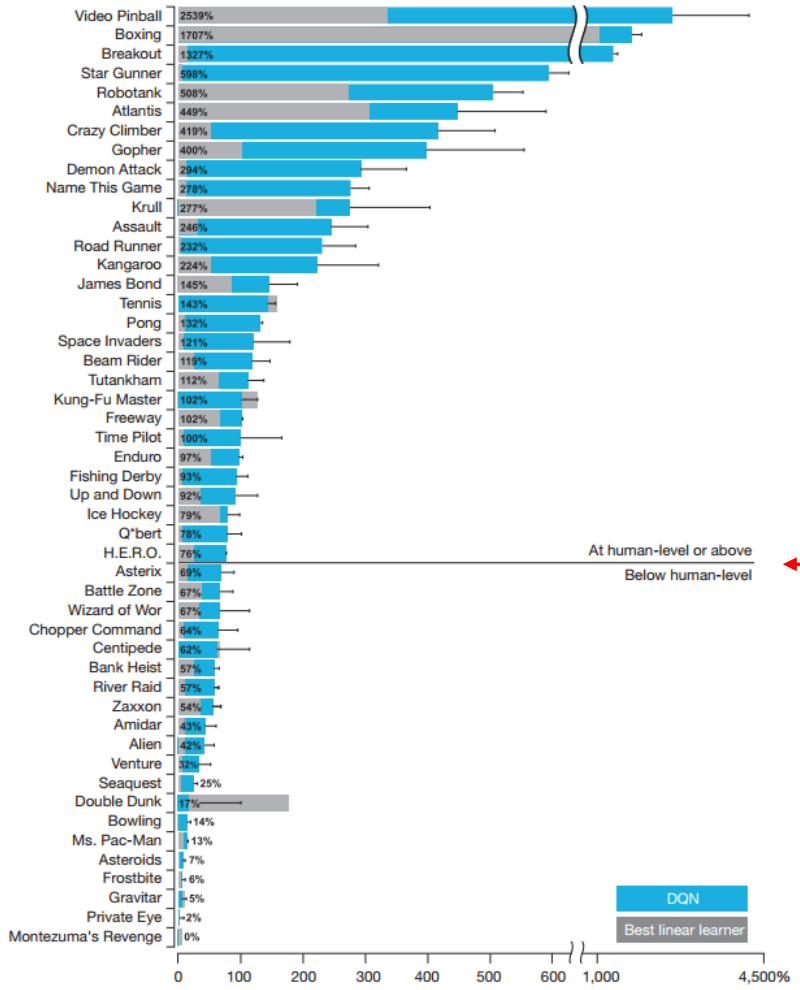
Input: 4 frames images (what you see on the screen) and score

Goal: maximize the score on the screen

No domain knowledge is involved!

The algorithm doesn't know what the control exactly do.

Atari Games



Tested on 49 games

Human-level

Mnih et al., 2015

Playing FPS Game (Lample et al., 2016)

https://www.youtube.com/watch?v=oo0TraGu6QY

Playing StarCraft (Peng et al., 2017)

http://v.youku.com/v_show/id_XMjcxMDgxNDg1Ng==.html

Playing StarCraft

http://v.youku.com/v_show/id_XMjcxMDgxNDg1Ng==.html

Playing StarCraft

http://v.youku.com/v_show/id_XMjcxMDgxNDg1Ng==.html

3 Conclusion

Conclusion

- Other applications:
  AlphaGo
  Robotic control
- The ability to tackle complex problems
- A way to artificial general intelligence (the intelligence of a machine that could successfully perform any intellectual task that a human being can)

Conclusion

References

Lample, G. and Chaplot, D.S., 2016. Playing FPS games with deep reinforcement learning. *arXiv preprint arXiv:1609.05521*.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G. and Petersen, S., 2015. Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.

Peng, P., Yuan, Q., Wen, Y., Yang, Y., Tang, Z., Long, H. and Wang, J., 2017. Multiagent Bidirectionally-Coordinated Nets for Learning to Play StarCraft Combat Games. arXiv preprint arXiv:1703.10069.


THANK
YOU!