Game Playing with DQfD and DQN

ADL Final Project

Team: Praise the Sun

Department of Computer Science, National Taiwan University

Abstract

Deep reinforcement learning are learning learning models that combines the traditional reinformance learning algorithms with modern state-of-the-art deep learning models. Currently, many applications, such as robotics and game playing, have reached astonishing performances with these kinds of models. However, these models require huge amount of time/self-generated data to achieve favorable results. What worse is that the performance is sometimes even unstable: the results may differ even with same environments and model settings. Some recent advanced algorithms may make the training process more efficient, for instance, the *Deep Q-Learning from Demonstrations*(DQfD) proposed by (T. Hester et al., 2017) [1], which introduced the so-called demonstration data that can train models in a supervised manner. In this project, we used some Atari games as the training environment, and we did some experiments on some traditional value-based deep reinforcement learning algorithms and on DQfD, and made some comparisons and analysis on our experiments.

Introduction

Aliquam non lacus dolor, *a aliquam quam* [**?**]. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nulla in nibh mauris. Donec vel ligula nisi, a lacinia arcu. Sed mi dui, malesuada vel consectetur et, egestas porta nisi. Sed eleifend pharetra dolor, et dapibus est vulputate eu. **Integer faucibus elementum felis vitae fringilla.** In hac habitasse platea dictumst. Duis tristique rutrum nisl, nec vulputate elit porta ut. Donec sodales sollicitudin turpis sed convallis. Etiam mauris ligula, blandit adipiscing condimentum eu, dapibus pellentesque risus.

Aliquam auctor, metus id ultrices porta, risus enim cursus sapien, quis iaculis sapien tortor sed odio. Mauris ante orci, euismod vitae tincidunt eu, porta ut neque. Aenean sapien est, viverra vel lacinia nec, venenatis eu nulla. Maecenas ut nunc nibh, et tempus libero. Aenean vitae risus ante. Pellentesque condimentum dui. Etiam sagittis purus non tellus tempor volutpat. Donec et dui non massa tristique adipiscing.

Main Objectives

1. Lorem ipsum dolor sit amet, consectetur.
2. Nullam at mi nisl. Vestibulum est purus, ultricies cursus volutpat sit amet, vestibulum eu.
3. Praesent tortor libero, vulputate quis elementum a, iaculis.
4. Phasellus a quam mauris, non varius mauris. Fusce tristique, enim tempor varius porta, elit purus commodo velit, pretium mattis ligula nisl nec ante.
5. Ut adipiscing accumsan sapien, sit amet pretium.
6. Estibulum est purus, ultricies cursus volutpat
7. Nullam at mi nisl. Vestibulum est purus, ultricies cursus volutpat sit amet, vestibulum eu.
8. Praesent tortor libero, vulputate quis elementum a, iaculis.

Materials and Methods

Fusce magna risus, molestie ut porttitor in, consectetur sed mi. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque consectetur blandit pellentesque. Sed odio justo, viverra nec porttitor vel, lacinia a nunc. Suspendisse pulvinar euismod arcu, sit amet accumsan enim fermentum quis. In id mauris ut dui feugiat egestas. Vestibulum ac turpis lacinia nisl commodo sagittis eget sit amet sapien.

Mathematical Section

Nulla vel nisl sed mauris auctor mollis non sed.

$$E = mc^2 \tag{1}$$

Curabitur mi sem, pulvinar quis aliquam rutrum. (1) edf (2) , $\Omega = [-1, 1]^3$, maecenas leo est, ornare at. $z = -1$ edf $z = 1$ sed interdum felis dapibus sem. x set y ytruem. Turpis j amet accumsan enim y-lacina; ref k-viverra nec porttitor x-lacina.

Vestibulum ac diam a odio tempus congue. Vivamus id enim nisi:

$$\begin{aligned}\cos\bar{\phi}_k Q_{j,k+1,t} + Q_{j,k+1,x} + \frac{\sin^2\bar{\phi}_k}{T\cos\bar{\phi}_k} Q_{j,k+1} = \\ -\cos\phi_k Q_{j,k,t} + Q_{j,k,x} - \frac{\sin^2\phi_k}{T\cos\phi_k} Q_{j,k}\end{aligned} \tag{2}$$

and

$$\begin{aligned}\cos\bar{\phi}_j Q_{j+1,k,t} + Q_{j+1,k,y} + \frac{\sin^2\bar{\phi}_j}{T\cos\bar{\phi}_j} Q_{j+1,k} = \\ -\cos\phi_j Q_{j,k,t} + Q_{j,k,y} - \frac{\sin^2\phi_j}{T\cos\phi_j} Q_{j,k}.\end{aligned} \tag{3}$$

Nulla sed arcu arcu. Duis et ante gravida orci venenatis tincidunt. Fusce vitae lacinia metus. Pellentesque habitant morbi. $\mathbf{A}\underline{\xi} = \underline{\beta}$ Vim $\underline{\xi}$ enum nidi $3(P + 2)^2$ lacina. Id feugain $\mathbf{A}$ nun quis; magno.

Results

Donec faucibus purus at tortor egestas eu fermentum dolor facilisis. Maecenas tempor dui eu neque fringilla rutrum. Mauris *lobortis* nisl accumsan. Aenean vitae risus ante. Phasellus imperdiet, tortor vitae congue bibendum, felis enim sagittis lorem, et volutpat ante orci sagittis mi. Morbi rutrum laoreet semper. Morbi accumsan enim nec tortor consectetur non commodo nisi sollicitudin. Proin sollicitudin. Pellentesque eget orci eros. Fusce ultricies, tellus et pellentesque fringilla, ante massa luctus libero, quis tristique purus urna nec nibh.

Nulla ut porttitor enim. Suspendisse venenatis dui eget eros gravida tempor. Mauris feugiat elit et augue placerat ultrices. Morbi accumsan enim nec tortor consectetur non commodo. Pellentesque condimentum dui. Etiam sagittis purus non tellus tempor volutpat. Donec et dui non massa tristique adipiscing. Quisque vestibulum eros eu. Phasellus imperdiet, tortor vitae congue bibendum, felis enim sagittis lorem, et volutpat ante orci sagittis mi. Morbi rutrum laoreet semper. Morbi accumsan enim nec tortor consectetur non commodo nisi sollicitudin.

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 1: Table caption

Figure 1: Figure caption

In hac habitasse platea dictumst. Etiam placerat, risus ac.
Adipiscing lectus in magna blandit:

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 2: Table caption

Vivamus sed nibh ac metus tristique tristique a vitae ante. Sed lobortis mi ut arcu fringilla et adipiscing ligula rutrum. Aenean turpis velit, placerat eget tincidunt nec, ornare in nisl. In placerat.

Figure 2: Figure caption

Conclusions

- Pellentesque eget orci eros. Fusce ultricies, tellus et pellentesque fringilla, ante massa luctus libero, quis tristique purus urna nec nibh. Phasellus fermentum rutrum elementum. Nam quis justo lectus.
- Vestibulum sem ante, hendrerit a gravida ac, blandit quis magna.
- Donec sem metus, facilisis at condimentum eget, vehicula ut massa. Morbi consequat, diam sed convallis tincidunt, arcu nunc.
- Nunc at convallis urna. isus ante. Pellentesque condimentum dui. Etiam sagittis purus non tellus tempor volutpat. Donec et dui non massa tristique adipiscing.

Forthcoming Research

Vivamus molestie, risus tempor vehicula mattis, libero arcu volutpat purus, sed blandit sem nibh eget turpis. Maecenas rutrum dui blandit lorem vulputate gravida. Praesent venenatis mi vel lorem tempor at varius diam sagittis. Nam eu leo id turpis interdum luctus a sed augue. Nam tellus.

References

[1] Todd Hester, Matej Vecerik, Olivier Pietquin, Marc Lanctot, Tom Schaul, Bilal Piot, Andrew Sendonaris, Gabriel Dulac-Arnold, Ian Osband, John Agapiou, Joel Z. Leibo, and Audrunas Gruslys. Learning from demonstrations for real world reinforcement learning. *CoRR*, abs/1704.03732, 2017.

[2] Niels Justesen, Philip Bontrager, Julian Togelius, and Sebastian Risi. Deep learning for video game playing. *CoRR*, abs/1708.07902, 2017.

[3] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin A. Riedmiller. Playing atari with deep reinforcement learning. *CoRR*, abs/1312.5602, 2013.

[4] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay. *CoRR*, abs/1511.05952, 2015.

[5] Hado van Hasselt, Arthur Guez, and David Silver. Deep reinforcement learning with double q-learning. *CoRR*, abs/1509.06461, 2015.

[6] Ziyu Wang, Nando de Freitas, and Marc Lanctot. Dueling network architectures for deep reinforcement learning. *CoRR*, abs/1511.06581, 2015.

Acknowledgements

Etiam fermentum, arcu ut gravida fringilla, dolor arcu laoreet justo, ut imperdiet urna arcu a arcu. Donec nec ante a dui tempus consectetur. Cras nisi turpis, dapibus sit amet mattis sed, laoreet.