Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Computer Science and Statistics

Utilizing Reinforcement Learning and Deep Neural Networks to teach Atari VideO Games to Play Themselves

Jack Cassidy

14320816

February 20, 2018

A Final Year Project submitted in partial fulfilment
of the requirements for the degree of BAI (Computer Engineering)

Declaration

I hereby declare that this project is entirely my own work and that it has not been submitted as an exercise for a degree at this or any other university.

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at `http://www.tcd.ie/calendar`.

I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at `http://tcd-ie.libguides.com/plagiarism/ready-steady-write`.

Signed: _____   Date: _____

Abstract

A short summary of the problem investigated, the approach taken and the key findings. This should be around 400 words, or less.

This should be on a separate page.

Acknowledgements

Thanks Mum!

You should acknowledge any help that you have received (for example from technical staff), or input provided by, for example, a company.

Contents

List of Figures

List of Tables

Nomenclature

A	Area of the wing	m^2
B		
C	Roman letters first, with capitals...	
a	then lower case.	
b		
c		
Γ	Followed by Greek capitals...	
α	then lower case greek symbols.	
β		
ϵ		
TLA	Finally, three letter acronyms and other abbreviations arranged alphabetically	

If a parameter has a typical unit that is used throughout your report, then it should be included here on the right hand side.

If you have a very mathematical report, then you may wish to divide the nomenclature list into functions and variables, and then sub- and super-scripts.

Note that Roman mathematical symbols are typically in a serif font in italics.

Introduction

Motivation

Machine Learning (ML) and Artificial Intelligence (AI) in 2018 are subjects that are almost unique in their ability to permeate into nearly every sphere, community and space in today's society. From the research community to the business world and the public eye through extensive media coverage, ML is certainly becoming more and more of a de facto part of our everyday lives.

Reinforcement Learning (RL) is a branch of ML that perhaps receives less public attention but is nonetheless believed to be set to revolutionize the field of AI (1). Recent breakthroughs in the application of Deep Learning to RL algorithms has spawned the exciting research field of Deep Reinforcement Learning (DRL) which has produced to date unparalleled results in various AI domains, such as defeating the world champion Go player (2).

There are is a vast quantity of RL methods and algorithms, such as Monte-Carlo, Q-Learning, SARSA, Policy Search and more. More recently, the advent of DRL has brought about adaptations to existing algorithms to expand their use to multi-dimensional observations spaces such as pixel information, such as Deep Q-Learning (3). It is easy to become overwhelmed with all of these offerings when exploring the RL space. The motivation behind this project is to demystify the state of the art of RL.

Objectives

The objectives of this project are threefold.

1. Research the history and state of the art of the RL domain.

2. Build a system on top of the Arcade Learning Environment (ALE) (4) to evaluate the performance of three state of the art DRL algorithms across a series of metrics when applied to a selection of Atari 2600 video games.

3. Carry out the experiments required to make these evaluations and draw conclusions from them.

The aim is for the system to strive towards generality. The system is given no prior knowledge of how each game works and there is no change in the underlying architecture of the solution when applied to different games, all while maintaining a high level of performance. The system achieved this goal across two games, Space Invaders (5) and Breakout (6), with sufficiently differing game mechanics (graphics, control mechanisms, scoring systems etc.) so as to somewhat qualify this claim. Due to the lengthy nature of training DL models, only two games were used.

Research Methods

Report Overview

Chapter 2 gives some necessary background information. It will discuss the current state of the art of RL with particular interest in how it is being applied to video games, as well as the technologies and tools being used in research today and for this project.

Chapter 3 outlines the architecture of the system and the rationale behind certain design decisions.

Chapter 4 will discuss the components of the experiment evaluation. It will give a greater elaboration of the project's objectives, a description of the experimental setup, and a discussion of the results.

Chapter 5 closes the project with a conclusion of all that has been discussed, an outline of what has been achieved from both an objective and personal point of view and finally a suggestion for future work.

Figures, Tables and Referencing

sIt is very important to properly refer in the text to any figures, tables or previously published work that you are discussing. Adequate and consistent referencing is one of the criteria which will be used to assess your project report.

Figures

Graphs, pictures and other images should be included in your report as a numbered, captioned figure. An example is given in Figure 2.1.
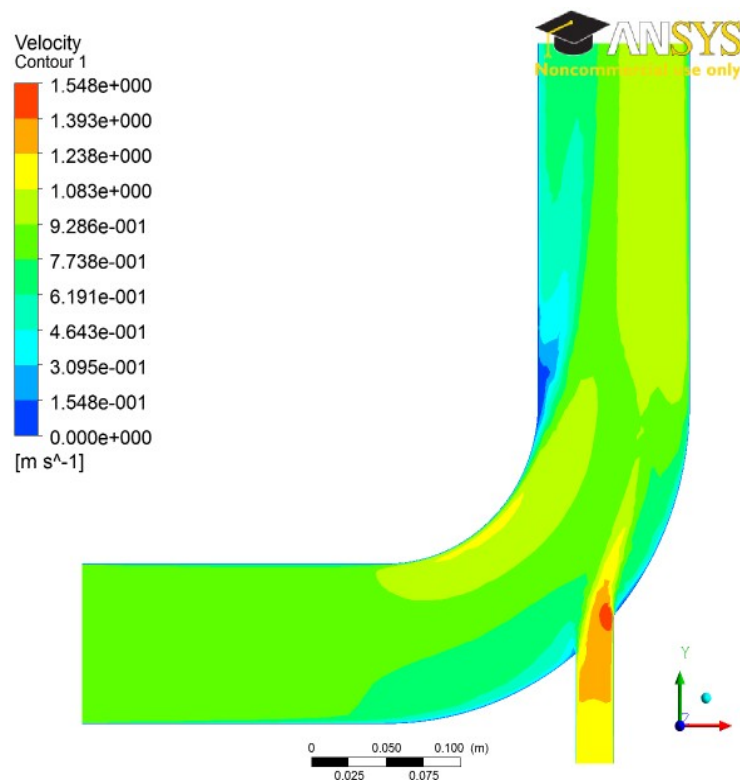

Figure 2.1: Velocity distribution on the mid-plane for an inlet velocity for case 1.

The figure and caption should be centred. The figure numbering starts at 1 at the beginning of each chapter. The caption should provide a brief description of what is being shown. The

figure should appear in the document after it is referred to in the text. No figure should be included which is not referred to in the text. Ensure that the size and resolution of images imported from software are sufficient to read any text.

Tables

Tables are an important way of displaying your results; Table 2.1 is a sample table, adapted from the Master/Doctoral Thesis template at `http://www.latextemplates.com/cat/theses`, which was generated with this code:

```
\begin{table}[b]
\caption{The effects of treatments X and Y on the four groups studied.}
\label{tab:treatments}
\centering
\begin{tabular}{l l l}
\toprule
\textbf{Groups} & \textbf{Treatment X} & \textbf{Treatment Y} \\\midrule
1 & 0.2 & 0.8\\
2 & 0.17 & 0.7\\
3 & 0.24 & 0.75\\
4 & 0.68 & 0.3\\
\bottomrule\\
\end{tabular}
\end{table}
```

Tables are numbered in the same way as figures. Typically tables also have a short caption, but this is not universally true. The number and caption appear above the table, not below as with figures. Again, no table should appear in the report which has not been referred to in the text. Tables should come after they are discussed in the text. The exact formatting of the table depends somewhat on the content of the table, but in general, the text in the table should be the same font and size as the main text.

Table 2.1: The effects of treatments X and Y on the four groups studied.

Groups	Treatment X	Treatment Y
1	0.2	0.8
2	0.17	0.7
3	0.24	0.75
4	0.68	0.3

Equations

All equations should be numbered sequentially. Do not restart the numbering at the beginning of each chapter. Unlike figures and tables, you may not need to refer to every equation in the text. You should take care to format equations properly. Do no simply try to use plain text. Use the equation layout facilities. An example of how equations should appear is shown in Equation 1. Here is the code for it:

```
\begin{equation}
\textrm{div}(\underline{u}) = \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} +
        \frac{\delta w}{\delta z} = 0
\label{sampleequation}
\end{equation}
```

$$\mathrm{div}(\underline{u}) = \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y} + \frac{\delta w}{\delta z} = 0 \tag{1}$$

Referencing published work

It is important to give appropriate credit to other people for the work that they have shared through publications. In fact, you must sign a declaration in your report stating that you understand the nature of plagiarism. As well as avoiding plagiarism, citing results or data from the literature can strengthen your argument, provide a favourable comparison for your results, or even demonstrate how superior your work is.

There are many styles to reference published work. For example, the parenthetical style (which is also called the Harvard style) uses the author and date of publication (e.g. "Smith and Jones, 2001"). There is also the Vancouver (or the citation sequence) style, which is shown in this document. In the Vancouver style, the publications are cited using a bracket number which refers to the list in the References section at the end of the report. The references are listed in order that they are cited in the report. A variant is name sequence style in which the publications are referenced by number, but the list is arranged alphabetically. For example, the text might say: several studies have examined the sound field around tandem cylinders generated by flow(7, 8), while other investigations have focused on the effect of an applied sound field on the flow(9). Papers from conference proceedings(10), books(11) and technical reports(12, 13) can be dealt with in the same style.

The Vancouver style has the advantage that it is a little more compact in the text and does not distract from the flow of the sentence if there are a lot of citations. However, it has the

disadvantage that it is not immediately clear to the reader what particular work has been referenced.

It actually does not matter which particular referencing style is used as long as three important considerations are observed:

- the referencing style used throughout the document is consistent;

- all material used or discussed in the text is properly cited;

- nothing is included in the reference list that has not been cited.

This template has a suitable referencing style already set up – you should use it and use the built-in BibTeX system to manage your references. See above for examples of how to cite a reference and look in the `sample.bib` file to see BibTeX references. Remember Google Scholar and other search engines will give you BibTeX references for lots of academic publications. Otherwise, you can easily make up your own based on the examples in that file.

LATEX

seeing LATEX, or more properly "LATEX 2_ε", is a very useful document processing program. It is very widely used, widely available, stable and free. Famously, TEX, upon which LATEX is built, was originally developed by the eminent American mathematician Donald Knuth because he was tired of ugly mathematics books(14). Although it has a learning curve (made much less forbidding by online tools and resources – see below), it allows the writer to concentrate more fully on the content, and takes care of most everything else.

While it can be used as a word processor, it is a *typesetting* system, and Knuth's idea was that it could be used to produce beautiful looking books:

> *LATEX is a macro package which enables authors to typeset and print their work at the highest typographical quality, using a predefined, professional layout.*[1]

LATEX has great facilities for setting out equations and a powerful and very widely supported bibliographic system called BibTeX, which takes the pain out of referencing.

Three useful online resources make LATEX much better:

(1) An excellent online LATEX environment called "Overleaf" is available at `http://www.overleaf.com` that runs in a modern web browser. It's got this template available – search for a TCD template. Overleaf can work in conjunction with Dropbox, Google Drive and, in beta, GitHub.

(2) Google Scholar, at `http://scholar.google.com`, provides BibTeX entries for most of the academic references it finds.

(3) An indispensable and very fine introduction to using LATEX called *"The not so short introduction to LATEX 2_ε"* by Oetiker et al. (15) is online at `https://doi.org/10.3929/ethz-a-004398225`. Browse it before you use LATEX for the first time and read it carefully when you get down to business.

Other tools worth mentioning include:

---

[1] This is from Oetiker et al. (15). Did we mention that you should minimise your use of footnotes?

- `Draw.io` – an online drawing package that can output PDFs to Google Drive – see `https://www.draw.io`.

Evaluation

Conclusion

Bibliography

[1] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. A brief survey of deep reinforcement learning. *CoRR*, abs/1708.05866, 2017. URL http://arxiv.org/abs/1708.05866.

[2] David Silver, Aja Huang, Christopher J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529:484–503, 2016. URL http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html.

[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. URL http://arxiv.org/abs/1312.5602.

[4] Marc G. Bellemare, Yavar Naddaf, Joel Veness, and Michael Bowling. The arcade learning environment: An evaluation platform for general agents. *CoRR*, abs/1207.4708, 2012. URL http://arxiv.org/abs/1207.4708.

[5] Space invaders wikipedia page. URL https://en.wikipedia.org/wiki/Space_Invaders.

[6] Breakout wikipedia page. URL https://en.wikipedia.org/wiki/Breakout_(video_game).

[7] JA Fitzpatrick. Flow/acoustic interactions of two cylinders in cross-flow. *Journal of Fluids and Structures*, 17(1):97–113, 2003.

[8] SL Finnegan, C Meskell, and S Ziada. Experimental investigation of the acoustic power around two tandem cylinders. *Journal of Pressure Vessel Technology*, 132(4):041306, 2010.

[9] JW Hall, S Ziada, and DS Weaver. Vortex-shedding from single and tandem cylinders in the presence of applied sound. *Journal of Fluids and Structures*, 18(6):741–758, 2003.

[10] Peter Jordan, John Fitzpatrick, and Craig Meskell. Array beam pattern control for measurement of propeller noise. In *AIAA/CEAS Aeroacoustics Conference and Exhibit, Maastricht, Netherlands*, 2001.

[11] Michael P Païdoussis, Stuart J Price, and Emmanuel De Langre. *Fluid-structure interactions: cross-flow-induced instabilities*. Cambridge University Press, 2010.

[12] L Reyes. Power uprate program status report-secy-07-0090. Technical report, Technical Report, US Nuclear Regulatory Commission, 2007.

[13] International Energy Agency. Technology Roadmap — Geothermal Heat and Power. `https://www.iea.org/publications/freepublications/publication/Geothermal_Roadmap.pdf`.

[14] Len Shustek and Donald Interviewee-Knuth. Interview donald knuth: A life's work interrupted. *Communications of the ACM*, 51(8):31–35, 2008.

[15] Tobias Oetiker, Hubert Partl, Irene Hyna, and Elisabeth Schlegl. The not so short introduction to latex 2ε. 2001. `https://doi.org/10.3929/ethz-a-004398225`.

Appendix

You may use appendices to include relevant background information, such as calibration certificates, derivations of key equations or presentation of a particular data reduction method. You should not use the appendices to dump large amounts of additional results or data which are not properly discussed. If these results are really relevant, then they should appear in the main body of the report.

Appendix numbering

Appendices are numbered sequentially, A1, A2, A3... The sections, figures and tables within appendices are numbered in the same way as in the main text. For example, the first figure in Appendix A1 would be Figure A1.1. Equations continue the numbering from the main text.