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 23, 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

1 Introduction **1**

1.1 Motivation 1

1.2 Objectives 2

1.3 Research Methods 2

1.4 Report Overview 2

2 State of the Art **4**

2.1 Scope of the State of the Art 4

2.1.1 Notable Moments in History of Machine Learning in Video Games . . 4

3 LaTeX **5**

4 Evaluation **7**

5 Conclusion **8**

A1 Appendix **10**

A1.1 Appendix numbering 10

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. Businesses employ recommender systems to suggest new products to their customers and predict the rise and fall of stock prices using function approximators like Deep Learning. Traditional home appliances are now outdated in favour of smarter, IoT systems that learn our habits and provide a more tailored experience.

ML is a broad umbrella term, encapsulating a variety of different approaches. Most ML tasks can be classified as either supervised or unsupervised learning. Deep Learning is fast becoming a popular and powerful technique in supervised learning, involving teaching artificial neural networks to approximate any function, given enough training data. Reinforcement Learning (RL), another subset of supervised learning, 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 (7).

There are a growing number of RL methods and algorithms, such as Monte-Carlo, Q-Learning, SARSA and Policy Search (8). 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, a notable example being 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 development and state of the art of RL.
2. Build a system to evaluate the performance of three state of the art DRL algorithms by collecting a series of metrics while applying each algorithm to a selection of Atari 2600 video games.
3. Carry out the experiments, obtaining values for game score, survival time and model loss. Compare and contrast the different algorithms using the metrics recorded.

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 aim for the system is to be a general solution, that it can be expanded to work for any number of games and algorithms in the future with ease of implementation. The algorithms used are Deep Q-Learning, Double Q-Learning and Dueling Q-Learning.

Research Methods

This project takes a *case study* based approach to the experimentation. The first phase of the project involves building the system to the specification outlined previously. The second phase treats each game entered into the system as an individual case study. The game ROM is given as input to the system. The game is simulated by a third party emulator of our choosing (discussed in chapter 3), from which the system extracts greyscale frames to learn from. The output of the system is an action that it has chosen to be optimal, selected from the discrete vector of possible actions as defined by the game's control scheme. This control scheme is not provided to the system, it determines it dynamically with each game. The action is fed back into the emulator and the cycle continues up to a terminating signal.

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.

State of the Art

Scope of the State of the Art

It would be impossible to write a state of the art covering the breadth of ML within the bounds of this report. Instead the scope will be refined to the state of ML and RL in video games today. Video games provide a useful testbed for quickly and efficiently testing general AI agents. In order to claim that an agent achieves general competency, it should be tested in a set of environments that provide a suitable amount of variation, are reflective of real world problems the agent might encounter, and that were created by an independant party to remove experimenter's bias (2). Although the application of ML generated AI agents to video games may seem novel, the end goal is not to produce agents for defeating world champion chess players, but to take these general agents and extend them to more pressing problems to humanity, of which there are endless possibilities.

Notable Moments in History of Machine Learning in Video Games

The first application of using computing to play a video game of notoriety arose in the research paper "Programming a Computer for Playing Chess" (5) as far back as 1950. In that paper, author Claude Shannon also highlighted the point that although the application of such a solution may seem unimportant;

> *"It is hoped that a satisfactory solution of this problem will act as a wedge in attacking other problems of a similar nature and of greater significance"*

Claude designed a strategy that, even at the time, was infeasible as it would take more than 16 minutes to make a move.
Fast forward to 1997, IBM came out with "Deep Blue," a computer purpose built to play chess. It is reknowned as the first AI system to defeat a world champion chess player, Garry Kasparov under normal game regulations.

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(6). 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. (4) 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. (4). 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] Arulkumaran, K., Deisenroth, M. P., Brundage, M., and Bharath, A. A. (2017). A brief survey of deep reinforcement learning. *CoRR*, abs/1708.05866.

[2] Bellemare, M. G., Naddaf, Y., Veness, J., and Bowling, M. (2012). The arcade learning environment: An evaluation platform for general agents. *CoRR*, abs/1207.4708.

[3] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. A. (2013). Playing atari with deep reinforcement learning. *CoRR*, abs/1312.5602.

[4] Oetiker, T., Partl, H., Hyna, I., and Schlegl, E. (2001). The not so short introduction to latex 2ε. `https://doi.org/10.3929/ethz-a-004398225`.

[5] Shannon, C. E. (1950). Xxii. programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(314):256–275.

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

[7] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 529:484–503.

[8] Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge.

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.