

Teach AI to play Atari iceHockey game using gameplay videos

Yangyang Xu

A report submitted for the course
COMP8755 ANU Individual Computing Project
Supervised by: Dr Penny Kyburz
The Australian National University

March 2019

© Yangyang Xu 2019

DRAFT - 25 March 2019

Except where otherwise indicated, this report is my own original work.

Yangyang Xu
25 March 2019

Acknowledgments

I would like to express my great appreciation to Dr. Penny Kyburz for her patient guidance and constructive suggestions. I also want to thank Dr. Penny Kyburz for offering me an opportunity to work on this interesting project.

Abstract

Replace the text here with your abstract.

This is an example of a \LaTeX report suitable for CS Projects. It has been adapted from a template ANU thesis and contains some examples of how to typeset code listings, figures and tables in a computer science context.

Contents

Acknowledgments	iii
Abstract	v
1 Introduction	1
1.1 Problem Statement	1
1.2 Motivations	1
1.2.1 Benefits of training AI to play games	1
1.2.2 Existing gap	1
1.3 Project Scope	1
1.4 Report Outline	1
2 Background and Related Work	3
2.1 Self-Supervised Learning	3
2.2 Markov Decision Processes (MDPs)	3
2.3 Inverse Reinforcement Learning (IRL)	3
2.4 Related work	3
2.5 Summary	3
3 Design and Implementation	5
3.1 Smart Design	5
3.2 Summary	5
4 Experimental Methodology	7
4.1 Software platform	7
4.2 Hardware platform	7
5 Results	9
5.1 Direct Cost	9
5.2 Summary	9
6 Conclusion	11
6.1 Future Work	11
Bibliography	13

List of Figures

4.1	Hello world in Java and C. This short caption is centered.	8
5.1	The cost of zero initialization	10

List of Tables

4.1	Processors used in our evaluation. Note that the caption for a table is at the top. Also note that a really long comment that wraps over the line ends up left-justified.	7
-----	---	---

Introduction

Put your introduction here. You could use `\fix{ABCDEFG.}` to leave your comments, see the box at the left side.

You have to
rewrite your
thesis!!!

1.1 Problem Statement

Describe the problem your project addresses.

1.2 Motivations

1.2.1 Benefits of training AI to play games

1.2.2 Existing gap

1.3 Project Scope

Describe the problem your project addresses.

1.4 Report Outline

How many chapters you have? You may have Chapter 2, Chapter 3, Chapter 4, Chapter 5, and Chapter 6.

Background and Related Work

At the beginning of each chapter, please give the motivation and high-level picture of the chapter. You also have to introduce the sections in the chapter, e.g.

Section ?? gives background material necessary in order to read this report ,and related work is given in Section 2.4.

2.1 Self-Supervised Learning

2.2 Markov Decision Processes (MDPs)

2.3 Inverse Reinforcement Learning (IRL)

2.4 Related work

You may reference other papers. For example: Generational garbage collection [Lieberman and Hewitt, 1983; Moon, 1984; Ungar, 1984] is perhaps the single most important advance in garbage collection since the first collectors were developed in the early 1960s. (doi: "doi" should just be the doi part, not the full URL, and it will be made to link to dx.doi.org and resolve. shortname: gives an optional short name for a conference like PLDI '08.)

2.5 Summary

Summarize what you discussed in this chapter, and introduce the story of next chapter. Readers should roughly understand what your report talks about by only reading words at the beginning and the end (Summary) of each chapter.

Design and Implementation

Same as the last chapter, give the motivation and the high-level picture of this chapter to readers, and introduce the sections in this chapter.

3.1 Smart Design

3.2 Summary

Same as the last chapter, summarize what you discussed in this chapter and be a bridge to next chapter.

Experimental Methodology

4.1 Software platform

We use Jikes RVM, which we defined in a macro in `macros.tex`. We ran the `avrora` benchmark, which we're typesetting in sans-serif font to make it clear it's a name.

You can also use inline code, like a `&& b`. Notice how, unlike when using the `texttt` command, the `icode` macro also scales the x-height of the monospace font correctly.

4.2 Hardware platform

Table 4.1 shows how to include tables and Figure 4.1 shows how to include codes. Notice how we can also use the `cleveref` package to insert references like Table 4.1, by writing just `\cref{tab:machines}`.

We can also refer to specific lines of code in code listings. The bug in Figure 4.1(a) is on line 4. There is also a bug in Figure 4.1(b) on lines 4 to 6. To achieve these references we put `(*@ \label{line:bug} @*)` in the code – the `(*@ @*)` are escape delimiters that allow you to add LaTeX in the (otherwise verbatim) code file.

Table 4.1: Processors used in our evaluation. Note that the caption for a table is at the top. Also note that a really long comment that wraps over the line ends up left-justified.

Architecture	Pentium 4	Atom D510	Sandy Bridge
Model	P4D 820	Atom D510	Core i7-2600
Technology	90nm	45nm	32nm
Clock	2.8GHz	1.66GHz	3.4GHz
Cores \times SMT	2 \times 2	2 \times 2	4 \times 2
L2 Cache	1MB \times 2	512KB \times 2	256KB \times 4
L3 Cache	none	none	8MB
Memory	1GB DDR2-400	2GB DDR2-800	4GB DDR3-1066

```
1  int main(void)
2  {
3      printf("Hello_World\n");
4      int a = *((volatile int *) 0); // uh-oh!
5      return 0;
6  }
```

(a) C

```
1  void main(String[] args)
2  {
3      System.out.println("Hello_World");
4      List<String> a = new ArrayList<String>();
5      a.add("foo");
6      List<Object> b = a; // This is a compile-time error
7  }
```

(b) Java

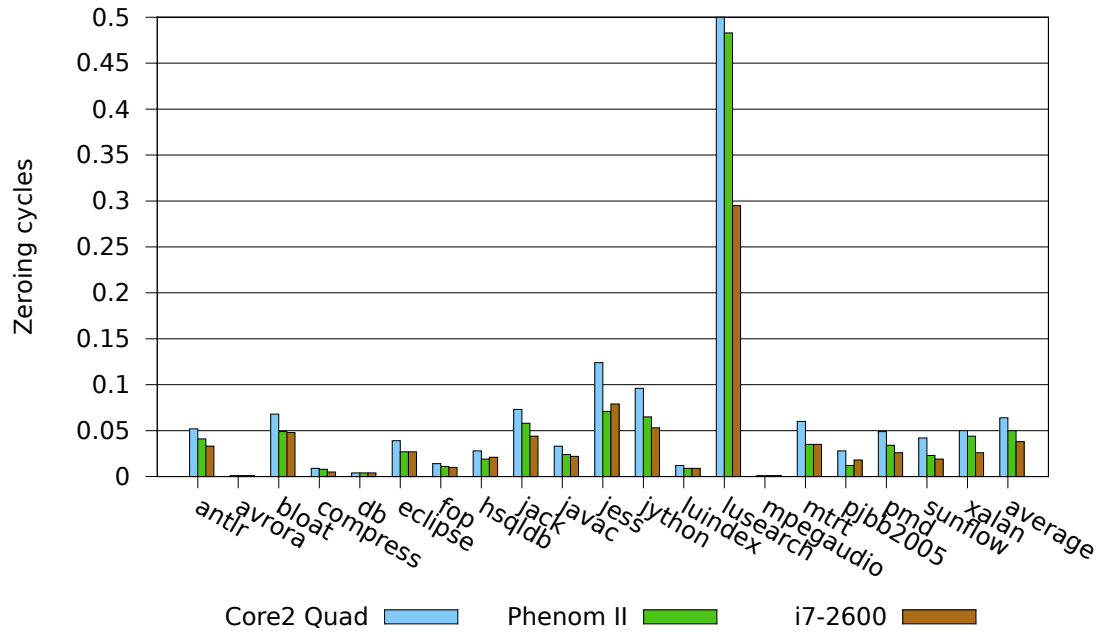
Figure 4.1: Hello world in Java and C. This short caption is centered.

Results

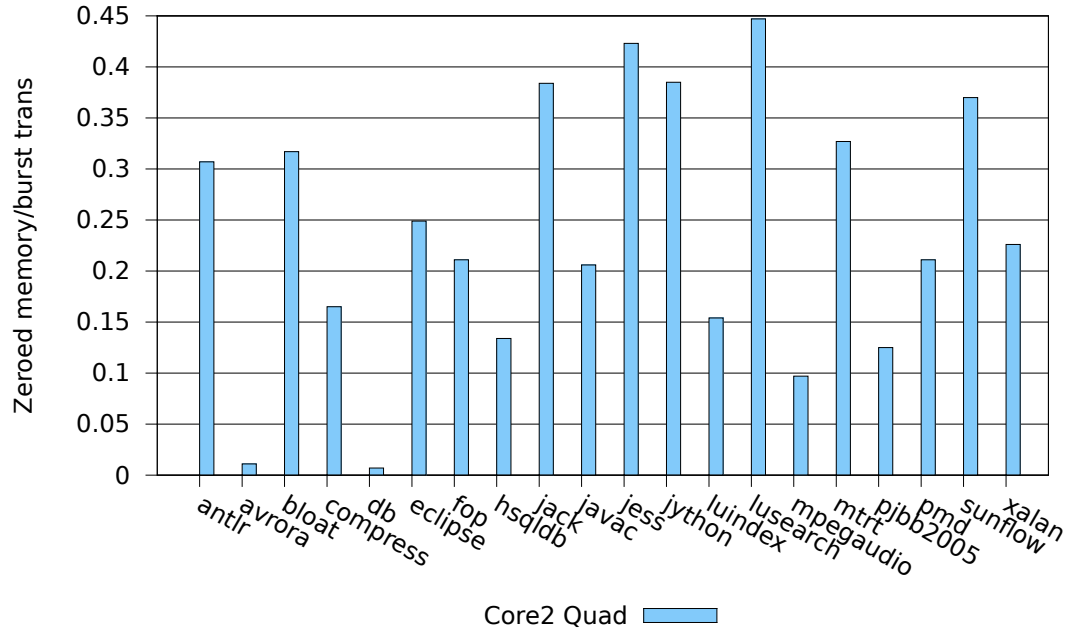
5.1 Direct Cost

Here is the example to show how to include a figure. Figure 5.1 includes two subfigures (Figure 5.1(a), and Figure 5.1(b));

5.2 Summary



(a) Fraction of cycles spent on zeroing



(b) BytesZeroed / BytesBurstTransactionsTransferred

Figure 5.1: The cost of zero initialization

Conclusion

Summarize your work, state your main findings and discuss what you are going to do in the future in Section 6.1.

6.1 Future Work

Good luck.

Bibliography

- LIEBERMAN, H. AND HEWITT, C., 1983. A real-time garbage collector based on the lifetimes of objects. *Communications of the ACM*, 26, 6 (Jun. 1983), 419–429. doi:10.1145/358141.358147. (cited on page 3)
- MOON, D. A., 1984. Garbage collection in a large LISP system. In *LFP '84: Proceedings of the 1984 ACM Symposium on LISP and Functional Programming* (Austin, Texas, USA, Aug. 1984), 235–246. ACM, New York, New York, USA. doi:10.1145/800055.802040. (cited on page 3)
- UNGAR, D., 1984. Generation scavenging: A non-disruptive high performance storage reclamation algorithm. In *SDE 1: Proceedings of the 1st ACM SIGSOFT/SIGPLAN Software Engineering Symposium on Practical Software Development Environments* (Pittsburgh, Pennsylvania, USA, Apr. 1984), 157–167. ACM, New York, New York, USA. doi:10.1145/800020.808261. (cited on page 3)