Research in Industrial Projects for Students

Sponsor

Advanced Micro Devices

Statement of Work

Exploration of Reinforcement Learning in Computer Games

Student Members

Jiajing Guan (Project Manager), *George Mason University*

Elvis Nunez, *Brown University*

Kaman Phamdo, *University of Maryland*

Patrik Gerber, *University of Oxford*

Academic Mentor

Tonmoy Monsoor, `mtonmoy@g.ucla.edu`

Sponsoring Mentors

Nicholas Malaya, `nicholas.malaya@amd.com`

Abhinav Vishnu, `abhinav.vishnu@amd.com`

Date: June 29, 2018

This project was jointly supported by Advanced Micro Devices and NSF Grant ID

1 Introduction

Advanced Micro Devices, Inc. (AMD) is an American multinational semiconductor company based in Santa Clara, California, that develops computer processors and related technologies for business and consumer markets. AMD's main products include microprocessors, motherboard chipsets, embedded processors and graphics processors for servers, workstations and personal computers, and embedded systems applications.

A major automotive company demoed an RC car at NIPS2017 that dynamically explored and interacted with its environment. As the car explored, it learned more about its environment and how to best navigate it. This is distinct from typical approaches in Artificial Intelligence, which operate in separate modes of training and inference. Often, an algorithm must first learn by viewing examples of a task being performed (training) before it is capable of making predictions for this same task (inference). The bifurcation between learning and action is distinct from naturally occurring intelligences, which acquire knowledge continuously through experience and trial-and-error. Mimicking the feedback loop between action and results is the motivation for the field of reinforcement learning, which provides a means for an agent to continually learn while dynamically interacting with an environment. Ultimately, the promise of reinforcement learning is the creation of general-purpose artificial intelligences that can interact with and learn from the world around them.

2 Problem Description

2.1 Problem Background

Typically, supervised machine learning approaches operate in distinct modes of training and inference. During training, an algorithm is shown labeled data and learns a representation of the underlying distribution of the data. After training, the model can be deployed to provide predictions in a phase known as inference. When such model is built upon layers of neurons interact with each other to simplify complicated concepts, it is called deep learning.

However, the field of reinforcement learning blends these approaches and provides a means for an agent to continually learn while dynamically interacting with an environment. A periodic reward provides the agent a feedback signal to guide it towards behavior that is considered beneficial. However, many complex tasks provide such feedback rarely and in latency, resulting in infrequent and inefficient training of the agent. As a result, even though high-dimensional sensory data such as vision has been an interest point for reinforcement learning, with the lack of hand-crafted featuring of the raw inputs, reinforcement learning algorithm hasn't been successful when given inputs like images.

Recent developments in deep learning have been successful in exacting features from high-dimensional sensory data, which make people wonder if such improvements can be used on reinforcement learning.

2.2 Problem Statement

This project will investigate modern approaches to provide a richer set of feedback and guide the agent towards more effective learning while exploring common frameworks and algorithms in reinforcement learning. The project will also investigate techniques to reduce the power consumption and latency of the deployed model without compromising the accuracy of the system. This will involve characterizing the impact of parameters such as the frame rate, learning rate and precision of numerical operations on the results of the problem, as well as optimizations for a particular hardware device. The results of this study will be applicable to edge inferencing autonomous vehicles, robotics, manufacturing, and smart grids.

3 Objectives

3.1 Main Objectives

The students will perform the following tasks:

- Characterize common reinforcement learning approaches, and relate their findings towards efficient implementations in hardware devices.
- Implement Q-learning on a simple computer game, such as the classic Atari games NeonRacer or Snakes

3.2 Additional Objectives

If time permits, the students will attempt the following objectives:

- Investigate techniques to reduce the power consumption and latency of the deployed model without compromising the accuracy of the system. This will involve characterizing the impact of parameters such as the frame rate, learning rate and precision of numerical operations on the results of the problem, as well as optimizations for a particular hardware device.
- Consider identifying algorithmic extensions to existing approaches. This will focus on applying and extending curriculum and imitation learning (providing simple examples for the agent to learn) and observing how successfully these approaches generalize.

4 Approach

The venue for this exploration will be computer games, which provide an ideal sandbox to explore reinforcement learning approaches.

The students will implement reinforcement learning algorithm while incorporate deep learning. The students will train the algorithm in the games provided by OpenAI gym as such games allow the algorithm to develop with relatively low cost. With a trained algorithm, the students will explore the effects of varying parameters in the environment. If time permitted, the students will provide a theoretical justification to the results obtained in the previous step.

5 Deliverables

- From Students to Sponsors

  - AMD site visit in Sunnyvale, CA (to be arranged)
  - Weekly conference call meetings
  - Midterm oral presentation
  - Midterm report
  - Projects Day presentation
  - Final report of findings
  - Software (if appropriate)
    * Specify sponsor-approved OS, platform Documentation
- From Sponsors to Students
  - Set up AMD site visit
  - Weekly conference call meetings
  - Project and platform specifications

6 Timetable

- Week 1 - 3
  - Read reinforcement learning and deep learning texts
  - Work on reproducing results from the paper *Playing Atari with Deep Reinforcement Learning* by Mnih., et al. by implementing Q-learning on a simple computer game
- Week 4 - 6
  - Analyze the impact of changes in various parameters on the performance and results of Q-learning model
  - Prepare and present midterm written report and oral presentation
- Week 7 - 9
  - Perform any further experiments based on previous results
  - Complete final report and Projects Day presentation

Reference

[1] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, *A brief survey of deep reinforcement learning*, arXiv preprint arXiv:1708.05866, (2017).

[2] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, *Openai gym*, arXiv preprint arXiv:1606.01540, (2016).

[3] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1, MIT press Cambridge, 2016.

[4] T. Matiisen, A. Oliver, T. Cohen, and J. Schulman, *Teacher-student curriculum learning*, arXiv preprint arXiv:1707.00183, (2017).

[5] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, *Playing atari with deep reinforcement learning*, arXiv preprint arXiv:1312.5602, (2013).

[6] Y. Tian, Q. Gong, W. Shang, Y. Wu, and C. L. Zitnick, *Elf: An extensive, lightweight and flexible research platform for real-time strategy games*, in Advances in Neural Information Processing Systems, 2017, pp. 2656–2666.