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

Introduction

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.

The sponsor of this project, 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. The results of this study will be applicable to edge inferencing autonomous vehicles, robotics, manufacturing, and smart grids.

Problem Description

Motivation

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 a model is built from layers of computational units called neurons that interact with each other, we call it deep learning.

The field of deep reinforcement learning blends techniques from the areas of deep and reinforcement learning to provide a means for an agent to continually learn while dynamically interacting with an environment. A periodic reward signal provides the agent feedback to guide it towards behavior that is considered beneficial. Many complex tasks provide such feedback rarely and with large latency, resulting in a difficult learning task. Because of this, classical reinforcement learning had serious limitations when applied to high-dimensional data. However, recent developments have shown deep reinforcement learning to be successful in some formerly intractable problems.

Problem Statement

This project will focus on improving the flexibility of reinforcement learning algorithms, allowing them to adapt to perturbations in the environment. The project will allow students to explore common frameworks and algorithms in reinforcement learning and investigate techniques to reduce the power consumption and latency of the deployed model without compromising the accuracy of the system. A possible avenue of exploration includes enriching the feedback received by the agent to guide it towards more effective learning. Moreover, analysis of model efficiency will revolve around characterizing the influence of hyperparameters, precision of numerical operations, and sensitivity to hardware devices.

Objectives

Main Objective

The focus will be to characterize common reinforcement learning approaches, and implement a state-of-the-art learning algorithm using techniques described in recent literature. The students will then test the algorithm on simple computer games, such as Snakes or the classic Atari game NeonRacer.

Additional Objectives

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

- Reduce the power consumption and latency of the learning algorithm without compromising the accuracy of the system. A possible approach to this includes analyzing the impact of parameters such as the frame rate, learning rate and precision of numerical operations on the algorithm's performance.

- Improve existing approaches by augmenting them with curriculum and imitation learning (providing simple examples for the agent to learn). Observe the robustness of this update and its flexibility.

Approach

Reinforcement learning agents learn by exploring their environment freely and receiving rewards for ideal behavior. A natural venue that provides such a sandbox is computer games. While employing deep learning methods, students will implement reinforcement learning algorithms to train agents to successfully solve computer games. Due to low cost in training, algorithms will be trained in the games provided by OpenAI Gym. Given a trained model, students will explore the effects of changing parameters in the environment; if time permits, students will provide a theoretical justification for these results.

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

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