1 Introduction

Nowadays a number of Fortune 500 companies are exploring opportunities to make use of Reinforcement Learning to resolve their industrial control problems. Now the major discussion in this concern is to figure out how to ameliorate Reinforcement Learning in order to solve their problems.

When we talk about Reinforcement Learning (RL) a lot of hype in this concern is created by solving the so-called Toy Problems. Toy Problems are available on open-source Reinforcement Learning tool-kits like Open AI Gym and Open AI Universe which serve as small piece of benchmark software providing a collection of test problems for developing and comparing reinforcement learning algorithms. Open AI gym makes no presumption about the structure of the agent, and is compatible with many numerical computation libraries, such as TensorFlow or PyTorch.

On the Contrary to Toy Problems, Enterprise customers face much more complex set of challenges when using Reinforcement Learning (RL) to control or optimize industrial applications.

Industrial control systems like Autonomous vehicles or Vertical takeoff, vertical landing for Rockets may involve dozens or thousands of variables, hence require human intensive calibration or optimization to generate reams of output data. Such control system is said to be high dimensional control systems. You might think that machine learning is utilized to solve complex problems all the time.

No, this is not always the cases. However, it is true that some problems can be best solved by supervised or unsupervised machine learning techniques. On the other hand, Reinforcement Learning (RL) which is an area of machine learning can help control and optimize some systems that other methods cannot.

As part of my thesis, I would deal with two control system Toy Problems using a set of the known techniques in Reinforcement Learning (RL). In order to tackle with these problems Deep Neural Networks (DNN) will be used as function approximators.

DNNs have been used in combination with Reinforcement Learning to exhibit impressive results for such high dimensional control systems. And this phenomenon of merging Deep Neural Networks with Reinforcement Learning is termed as Deep Reinforcement Learning.

This chapter will present the context for our work (Section 1.1). Then We follow it with problem definition that establishes the motivation behind the Thesis (Section 1.2), (Section 1.3) will describe the

research methodology we will follow and (Section 1.4) also describes the goals of our work. At the end we would provide the structure for the subsequent chapters (Section 1.5).

1.1 Context for our work

1.1.1 Historical Inspiration

From the time of the Industrial Revolution (1750-1850), there has been a demand to automate manual solutions which led to the Machine Revolution (1870-1940) replacing some of the work done physically by Machines. Due to the Machine Revolution masses of productivity increase were noticed worldwide and this fed to a new stage called Information Age(1950-now). Information Age was all about automating mental solutions. As a classical example performing difficult mental calculations with help of computers, which can automate the process of calculation in a way which very fast and very precise and hence replace the slower version of mental calculation performed by Humans. As a matter of fact, Information Age also led to a lot of productivity increase. But these phases had something in common that we had to invent a new solution every time in order to either automate physical or mental problems. Contrary to these phases there is another domain that is already ongoing and that allows machines to find solutions by themselves which are currently known by the name of Artificial intelligence (AI). Artificial Intelligence has huge potential upside cause if machines can determine solutions themselves then it takes away the responsibilities of Humans to find solutions for a particular problem. And as Humans, our task would be just to specify a problem and our goals and let the machine figure out how to solve that problem. In order for this to happen the Machine has to be Artificially made extensively intelligent so that it can autonomously learn to make decisions by itself. Nonetheless Having a huge potential upside and currently being an in-demand field of research, Artificial Intelligence has been on top of the investigation for many decades now. One of the first papers which put some light on whether a machine can attain decision-making capabilities and purposed to answer the question “Can machines think?” was published by Alan Turing and was titled COMPUTING MACHINERY AND INTELLIGENCE -1950. It quotes: -

“In the process of trying to imitate an adult human mind we are bound to think a good deal about the process which has brought it to the state that it is in. We may notice three components,

1. The initial state of the mind, say at birth,
2. The education to which it has been subjected,
3. Other experience, not to be described as education, to which it has been subjected.

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child’s? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child-brain is something like a note-book as one buys it from the stationers. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed “ – Alan Turing, 1950

In this quote, Turing is essentially talking about the process of learning and elaborates that it could be difficult to actually write a program that can replicate an adult mind rather he is conjecturing that it could easy to write a program that could learn by itself as a child does by interacting with the world and gaining experience from it.

1.1.2 Artificial Intelligence and Reinforcement Learning

There are numerous definitions of Artificial Intelligence termed by scientists and researchers over the decades. However, to be capable of making accurate decisions in order to accomplish certain goals is the definition of Artificial Intelligence used in this Thesis. And as per this definition of Artificial Intelligence: learning, decision-making and accomplishing goals are of prime focus. This then brings us to the definition of Reinforcement Learning which Alan Turing described in the quote above that People and Animals Agents learn by interacting with the environment.

And these Interactions are often sequential and future interactions might depend on earlier ones. The main purpose of these interactions is to accomplish a specific goal that certainly can be feasible after learning from the environment.

1.1.3 Relevance of Reinforcement Learning in the industry

Reinforcement Learning (RL) is one of the most exciting fields of Machine Learning today and also the oldest.

With an estimated market size of 7.35 billion US dollars, Artificial Intelligence is growing by leaps and bounds. McKinsey predicts that AI techniques (including Deep Learning and Reinforcement Learning) have the potential to create between 3.5T dollars and 5.8T dollars in value annually across nine business functions in 19 industries.

As explained in the (Section 1.1.1) , RL has been around since the 1950s, producing many interesting applications over the years, particularly in games (e.g., TD-Gammon, a Backgammon-playing program) and in machine control, but seldom making the headline news. But a revolution took place in 2013, when researchers from a British startup called DeepMind demonstrated a system that could learn to play just about any Atari game from scratch, eventually outperforming humans in most of them, using only raw pixels as inputs and without any prior knowledge of the rules of the games.

This was the first of a series of amazing feats, culminating in March 2016 with the victory of their system AlphaGo against Lee Sedol, a legendary professional player of the game of Go, and in May 2017 against Ke Jie, the world champion. No program had ever come close to beating a master of this game, let alone the world champion. Today the whole field of RL is boiling with new ideas, with a wide range of applications. DeepMind was bought by Google for over 500 million dollars in 2014.

1.2 Problem Definition

Now over the years number of Deep Reinforcement learning algorithms have been created but there are a number of challenges that one faces as consequences of these establishments as well.

These include :-

1. How to apply these algorithms to constrained Environment conditions?
2. Analyze which algorithm is better in a particular environment?
3. What is the Agents behavior?

These unaddressed consequences laid the foundation of the research , and these are the questions the research will try to answer.

1.3 Research Method

In this Thesis, we use the CRISP-DM [WH00] process model as our research methodology.

This is a chosen method for building models based on data mining, since agents and

their configurations are somehow also models on how a learning process should occur,

such that another learning model (i.e, the neural network, or brain of the agent) is

properly built.

The steps of the CRISP-DM methodology are shown in Figure 2.11.

• Business understanding

In this phase we need to understand and answer to the following questions:

What are representative engineering applications and what is the role of reinforcement learning in engineering problems?

How to build a DRL model according to specific problems?

Chapter 3 records our efforts in answering these research questions.

• Data understanding

In this phase we seek to understand the ideas behind the DQN, DDPG, PPO

methods we compare in this paper and understand the environment data selected

for our study.

Chapter 2 and Chapter 4 collect our results from this phase.

• Data preparation

In this phase we define the reward, observation and action space of the environments

and the learning steps of the training phase.

The results for this phase are summarized in Chapter 4.

• Modeling

Here we prepare and program the DRL agents and define the neural network

structures.

The results for this phase are summarized in Chapter 4.

• Evaluation

We compare the performance of the three methods according to the research

questions and evaluate them regarding the evaluation factors.

The results for this phase are summarized in Chapter 5.

1.4 Aim of the Research

We compare the performance of the three methods according to the research

questions and evaluate them regarding the evaluation factors.

The results for this phase are summarized in Chapter 5.