Imperial College London

Department of Computing

Data Efficient Deep Reinforcement Learning using Inductive Logic Programming

Author:
Kiyohito Kunii

Supervisor:
Professor Alessandra Russo

Submitted in partial fulfillment of the requirements for the MSc degree in MSc in Computing Science of Imperial College London

June 2018

Contents

Chapter 1

Introduction

There has been successful applications of deep reinforcement learning (DRL) a number of domains, such as games and robotics. DRL is considered to be a step towards artificial general intelligence. However there are still a number of issues to overcome in this method. As pointed by XXX, there are 3 major problems. First, it requires a large amount of data for training the model, which requires a long time of learning process. Second, it is considered to be a black-box, meaning the decision making process is unknown to human and therefore not explanable. Third there is no thoughts process to the decision making, which, as XX points out, is a fundamental to the artificial general intelligence. To overcome these problems, there are 3 main streams of research on this field. First main research is focused on applying Baysian statistics XXX, the second main research is XXX. Recently, the XXX attempted to incorporate symbolic representations into the system to achive more data-efficient learning, which shows a promissing results of this approach. The paper was the application of symbolic representatiojns into a very simple game to demonstrate this proof of concept would actually work.
By contract, there is active reserach in symbolic machine learning, which focuses on logic=based learning rather than statistical machine learning. For example, XX shows the agent can learn XXX from a noisy examples, with only a very few training examples.
In this paper, I explore incorporation of symbolic machine learning into XXX to achieve data-efficient learning using Inductive Learning of Answer Set Programs (ILASP), which is the state-of-art symbolic learning method that can be applied to imcomplete and more complex environment.
This research is based on XX, but in this paper I explore symbolic represetations process and include more learning aspect.
Problem setting Garnelo et al. (2016)

Chapter 2

Background

2.1 Data Efficient Learning

Two most studied approach for using previous learning exprience is meta-learning and transfer learning
Artificial General Intelligence
Definision of AGI
The history of data-efficient learning What other people have done in this.
The advance of statistical machine learning methods, especially deeep reinforcement learning
AlphaGo, and AlphaGo Zero
Also business success.
Study of symbolic machine learning roots from
Baysian Optimisation
RNN approach
Symbolic Deep reinforcement learning
Some implementation: German paper

2.2 ASP in Reinforcement Learning

Chapter 3

Background

3.1 Inductive Logic Programming

Inductive Logic Programming (ILP) is a subfield of research area aimed at the intersection of machine learning and logic programming (MUGGLETON 1991). The purpose of ILP is to derive a hypothesis H that is a solution of a learning task, which coveres all positive examples and any of negative examples.
Herbrand Model
Least Herbrand Model
Definite Logic Program is a set of definite rule where
a definite rule is of the form h $\leftarrow a_1, ..., a_n, whereh, a_1, ..., a_n areallatoms.$
whereas
Normal Logic Program is a set of normal rule where
a normal rule is of the form h $\leftarrow a_1, ..., a_n, notb_1, ..., notb_n wherehistheheadoftherule, anda_1, ..., a_n, b_1,$

3.1.1 Stable Model Semantics

Stable Model of a normal logic program

3.1.2 Anwer Set Programming

Literals
Negation as a failure
constrains is of the form $\leftarrow a_1, ..., a_n, notb_1, ..., notb_n$
Constraints are to filtering any irrelevant answer sets.
Syntax Examples
There are two types of constraints: soft and hard constraints. Soft constraint is XXX
Hard constraint is XXX
optimisation statement is of the form.
Which is useful to order the answer sets in terms of preference.
Syntax examples.
Aggregate
choice rules

An ASP program P is normal logic program with addition of choice rules, constraints and optimisation statement.
Answer set of P is

Cautious Induction

Cautious inductiion is based on the cautious semantics of answer set programs.
Sakama 2008
no concept of negative examples.

Brave Induction

Brave Induction is
Brave indction cannot learn constrains
Sakama 2009

3.1.3 Learning from Anwer Sets (LAS)

Learning from Answer Sets was developed in XX to overcome limitations of cautious induction and brave induction.
Partial Interpretations of the form

3.1.4 Inductive Learning of Answer Set Programs (ILASP)

ILASP is an algorithm that is capable of solving LAS tasks
It is based on two fundamental concepts: positive solutions and violating solutions.
A hypothesis H is a positive solution if and only if 1. H $\subseteq S_{\mathrm{M}}$
$2.\forall e^{+} \in \exists A \in AS(B \cup H) subjecttoAextendse^{+}$

A hypothesis H is a violating solution if and only if 1. H $\subseteq S_{\mathrm{M}}$
$2.\forall e^{+} \in E^{+} \exists A \in AS(B \cup H) subjecttoAextendse^{+}$
$3.\exists e^{-} \in E^{-} \exists A \in AS(B \cup H) subjecttoAextendse^{-}$

$\mathrm{ILP}_{\mathrm{LAS}}$ is positive solutions that are not violating solutions.
ILASP task containing a contex-dependent example
TODO Explain how symbolic learning works
TODO What would you learn in my context? Relashinship of the objects? Objects, types, locations and interactions.

3.2 Reinforcement Learning

An RL agent may include either Policy, Value function or Model,
where

3.2.1 Markov Decision Process(MDP)

3.2.2 Temporal-Difference Learning

3.2.3 Optimal Control Threads

3.2.4 Q-Learning

Bellman Equation

3.3 Symbolic Reinforcement Learning

Explain the paper

3.4 GVGAI Framework

Chapter 4

Project Overview

4.1 Motivation

Why symbolic reinforcement learning is good attempt
Reason 1: Complehensive by humans -¿ Explanable rather than black-box Reason 2: Similar to the human, it uses reasoning
Use of previous experience (background knowledge)
Not much explored yet.
However there is a room for exploration on this field.
TODO Explain how reinforcement learning works
Reason 3: Recent advance of ILASP is promissing
Because of the recent advancement of logic-based learning and deep reinforcement learning, combination of both approach would be a next explonation toward artificial general intelligence.

4.2 Objectives

combining the two novel approaches to overcome the problems of the QND

4.3 Project outline

The probject outline
Implement baseline performance (DQN)
Pipeline Implement the CNN side that is able to extract features of the game and convert into ASP syntax.
Apply ILASP to the ASP, which involves development of the pipeline of ILASP in Python
Finally use Q-learning that allow
which measurement would you use? (grid word, something else? GVGAL games)
Summarise different types of knowledge representations (Objects ?? relationship?)
Common sense
Implement based on T̈owards Deep Symbolic Reinforcement Learning

4.4 Contribution

4.5 Legal and Ethical Issues

???

4.6 References

Bibliography

Garnelo, M., Arulkumaran, K., and Shanahan, M. (2016). Towards deep symbolic reinforcement learning. *arXiv preprint arXiv:1609.05518*. pages 1