Université libre de Bruxelles

Analyzing MARL algorithms in dynamic environments: Evaluating performance with an additional unknown element

Preparatory work for the master thesis -- MEMO-F-403

Promoter:
Yannick MOLINGHEN

Author:
Kevin VANDERVAEREN

Supervisor:
Prof. Tom LENAERTS

Abstract

#todo{Abstract}

Table of Contents

I Introduction

I.1 Background and Objectives

A Multi-Agent Reinforcement Learning (MARL) is a subfield of the Reinforcement Learning domain which focuses on the interaction between multiple agents in a shared environment. Through the recent years, an increasingly amount of research has been conducted in this field to resolve issue that has arisen in the reel world. However, most of the research are done through simulations on environments which does not involve unknown elements in existing envirement. This thesis aims to evaluate the learning performance of MARL algorithms from a know environment to an slightly modified one by adding an unknown elements. Under the suppervision of Prof. Tom Lenaerts, and advisor Yannick Molinghen, from the Machine Learning Group (MLG) of the Université Libre de Bruxelles (ULB).

Currently, the research is focused on the the environment of LLE (Laser Learning Reinforcement) which is a environment created by Yannick Molinghen based on the original game . The environment is a 2D world also known as grid world where a single or multiple agents will be interacting. The goal of the agent is to reach an exit point while aquiring rewards (under the form of Gems) and avoiding obstacles.

The objective of the Master thesis is to develop a new feature in the LLE environment that was also includes in the original game of Oxen. Morever this feature has also a other objective which is to add a new element in the environment which are not included in the agents learning process and thus reevaluate the performance of a already fine tuned model that is trained on the original environment.

I.2 Notations and Definitions

Notations	Description
a	an action
s	

"T", "the transition function"

(temporary place) temporary place for some definition
- State Space S
- Agent i action space A_i
- joint Action Space $A = A_1 \times A_2 \times ... \times A_n$ where n is the number of agents
- specific action in the joint action space denoted a st. $a \in A$
- transition function T
- reward function R

the notation will be using the notation of ref to book here

(end temporary place)

II State of the Art

II.1 Distributed artificial intelligence

!!(this section has content form the article "Cooperative Multi-Agent Learning The State of the Art" by [ref to article])!!

Distributed artificial intelligence (DAI) is the a field of study which is rising in the last two decades. which is mainly focused on the domain of distributed systems. A distributed system by the definition of [ref to book] is a (quote)"where a number of entities work together to cooperatively solve problems"(endquote). this kind of study is not new, it has been studied for a long time. But what is new is the rise of the internet and the multiple electronic devices that we have today. Which bring the need of a new field of study which is the DAI that simply is the study of the interaction between multiple artificial intelligence (AI) or agents in a distributed system.

II.1.1 Multi-Agent Systems vs. Distributed problems Solving

In the field of DAI, we can find two main subfields a more traditional one which is the Distributed Problem Solving (DPS) which us the paradigm of a divide and conquer. The DPS is a field which is focused on distributing the problem to independent slaves which are solving the problem independently. On the other hand, the Multi-Agent Systems (MAS) emphasizes on the interaction between the agents.

II.1.2 Multi-Agent Systems

In MAS there are few constraints that are imposed on the agents. such as even though the agents are working together to solve a problem in a same environment they are not able to share their knowledge of the envirement with each other they may only acces to the information that they have, in RL we often refer this as a local obsevation. This is a important point because if they were able to share their knowledge this would be able to simply syncronize their knowledge and solve this problem as a DPS problem if the problem need no interaction between the agents (`todo` may be more).

II.2 Multi-Agent Learning

The Multi-Agent Learning (MAL)

II.3 Machine Learning

II.4 Single Agent Reinforcement Learning

II.4.1 Markov Decision Process

The Markov Decision Process (MDP) is a model that is used to often represented as a 5-tuple $\langle S, A, T, R \rangle$ where the elements are:

- S is the state space which can be represended as a all the possible configurations of the environment

II.4.1.1 State

A ways to represent the environment is to use a state. A state is an abstract ways to decribe the joint information of all elements in the environment. we can use as exemple the game of tick-tac-toe where the representation of the board at a given time such as this image [im] is a state. But a state is not only the representation of the board but also the information of the player turn. So a state is a representation of the environment at a given time. In the mathematical notation we usually use the notation s to represent a state, and S to represent the state space. The state space is the set of all possible states imagineable for a given environment.

II.4.2 Action

A Action reffers to the possible movement doable by the agent in the environment. In the case of the game of tick-tac-toe, the possible actions is to put a mark in one of the 9 squares. In the mathematical notation we usually use the notation a to represent an action, and A to represent the action space.

II.4.3 Transition

The transition is the function that is used to represent the change of a given state

II.5 Multi-Agent Reinforcement Learning

II.5.1 Stationary vs. Non-stationary

The MARL can be naively seen as adding more than