FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

František Dostál

Fitness and novelty in evolutionary reinforcement learning

Name of the department

Supervisor of the master thesis:	Mgr. Roman Neruda, CSc.
Study programme:	Mgr.
Study branch:	Artificial Inteligence

Prague 2023

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

Dedication.

Title: Fitness and novelty in evolutionary reinforcement learning

Author: František Dostál

Katedra teoretické informatiky a matematické logiky: Name of the department

Supervisor: Mgr. Roman Neruda, CSc., department

Abstract: Novelty is a novel aproach to modeling selection criteria in evolutionary algorithms and has been proven as viable technique of avoiding pitfalls of false optima in tasks abundant with them, such as solving mazes. Rather than closing the topic however, this finding opened other problems to explore: How to properly apply novelty in tasks that yield slightely better to conventional aproaches? How to properly model behavioral space nessessary for novelty computation? In this thesis we investigate use of novelty in selected reinforcement learnig tasks, it's combinations with classical fitness and propose behavior space models for the respective RL tasks.

Keywords: evolution novelty fitness behavioral space

Contents

Introduction 2

1 Introduction to Artificial Agents 3
- 1.1 Reactive agents . . . 3
- 1.2 Reinforcement learning . . . 3

2 Evolutionary algorithms 4
- 2.1 • . . . 4
- 2.2 Numerical Algorithms . . . 4
  - 2.2.1 Evolutionary Strategies . . . 4
  - 2.2.2 Differential Evolution . . . 4
- 2.3 Utility functions . . . 4
  - 2.3.1 Fitness . . . 4
  - 2.3.2 Pure Novelty . . . 4
  - 2.3.3 Combinations of Novelty . . . 4

3 Enviroments 5
- 3.1 Gymnasium . . . 5
  - 3.1.1 Cartpole . . . 5
  - 3.1.2 Lunar Lander . . . 5

4 Comparison criteria 6

5 Experiments 7

Conclusion 8

Bibliography 9

List of Figures 10

List of Tables 11

List of Abbreviations 12

A Attachments 13
- A.1 First Attachment . . . 13

Introduction

1. Introduction to Artificial Agents

To understand the reasoning guiding this work we have to first look at the fundamental subject of our efforts - the artificial agents.

Definition 1. *[Russell and Norvig, 2021] A rational agent is one that acts so as to achieve best outcome or, when there is uncertainty, the best expected outcome.*

1.1 Reactive agents

Definition 2 (Russell and Norvig [2021])**.** *A rational agent is one that acts so as to achieve best outcome or, when there is uncertainty, the best expected outcome.*

Definition 3. *A reward is a numerical evaluation of agent's behavior within a given enviroment. A reward function then is a function which assigns pair of state and action a real numerical value.*

1.2 Reinforcement learning

2. Evolutionary algorithms

2.1 •

2.2 Numerical Algorithms

In this section we will be reviewing different aproaches to implementation of evolutionary algoAlgorithmsrithms which were considered for the purpose of this thesis. Because we are optimising The algorithms were selected based on usefulness for the purposes of numerical optimalisation but also for their representativness of certain aspects of the field. Since the aim of this work is to compare aproaches of formulating utility function optimised by the evolution rather than the algorithms as such, we will consider them in their simplest form still expedient for satisfactory optimalisation of selected tasks.

2.2.1 Evolutionary Strategies

Evolutionary strategies are

2.2.2 Differential Evolution

2.3 Utility functions

2.3.1 Fitness

2.3.2 Pure Novelty

2.3.3 Combinations of Novelty

3. Enviroments

Since computer vision tasks are its own research domain, not yet fully explored, we conclude that tackling difficulties that would reliance on coputer vision model bring is not within the scope of this thesis. Consequently we focus mainly on enviroments with more semantically rich output.

3.1 Gymnasium

3.1.1 Cartpole

Cartpole is one of the most known benchmark enviroments in the Gymnasium library. It's a simple enviroment where the agent controls a cart rolling on flat surface with a pole on top. The goal is to keep the pole balanced on top of the cart. The reward is calculated every step dpending on the angle

3.1.2 Lunar Lander

Lunar Lander is somewhat more complex domain to optimise for than the Cartpole. Every step the reward is calculated based on several criteria, as oposed to only an angle of a stick, controling the landers descend and additional points are awarded for completing certain tasks. Additionally there are several variables determining the enviroments behaviour during evaluations such as the strenght of wind Unlike in cartpole where

4. Comparison criteria

5. Experiments

Conclusion

Bibliography

Stuart Russell and Peter Norvig. *Artificial Intelligence, Global Edition A Modern Approach*. Pearson Deutschland, 2021. ISBN 9781292401133. URL `https://elibrary.pearson.de/book/99.150005/9781292401171`.

List of Figures

List of Tables

List of Abbreviations

A. Attachments

A.1 First Attachment