

# Hindsight Goal Prioritization for Sparse Reward Environments

Final Project

Reinforcement Learning

**Flavio Maiorana** (2051396)

08/09/2023



SAPIENZA  
UNIVERSITÀ DI ROMA



# Table of Contents

## 1 Introduction

► Introduction

► Algorithm comparison

► Implementation

► Results



# Problem statement

## 1 Introduction

### Robotics environments

- Complex and different goals
- Sparse rewards
- Continuous action space



### Problems with exploration and reward shaping

- Goal may be too complex and observation space is big: we may never get reward 1
- Classical off-policy algorithms don't valorize much the failed episodes



Solution: Enhancing the Replay Buffer



# Fetch

## 1 Introduction

- Based on the 7-DoF Fetch Manipulator arm, with a two-fingered parallel gripper
- Tasks: Reach, Push, Slide and Pick-and-Place
- Action: Box(-1.0, 1.0, (4,)), float32)  $\rightsquigarrow$  Displacement in meters of the EE
- Observation: dictionary with info about the robot's end effector state and goal
  - Observation: ndarray of shape (25,)  $\rightsquigarrow$  kinematic info of the block object and EE
  - Desired goal: ndarray of shape (3,)  $\rightsquigarrow$  desired position of the EE or the block
  - Achieved goal: ndarray of shape (3,)  $\rightsquigarrow$  current position of the EE or the block
- Reward: if we use sparse rewards -1 for every timestep and 0 for reaching the goal
- Termination: episodes have no termination since they have infinite horizon. Thus, they are truncated after T steps (by default 50)



# Table of Contents

## 2 Algorithm comparison

► Introduction

► **Algorithm comparison**

► Implementation

► Results



# DDPG

The base of all

## Intuition

```
\documentclass{beamer}
```

## Drawbacks

```
\documentclass{beamer}
```



# HER

Sampling future goals

To start working with sintefbeamer

## Intuition

```
\documentclass{beamer}
```

## Drawbacks

```
\documentclass{beamer}
```



# HGR

Prioritizing future goals

## Intuition

```
\documentclass{beamer}
```

## Enhancements over Vanilla HER

```
\documentclass{beamer}
```





# Table of Contents

## 3 Implementation

► Introduction

► Algorithm comparison

► **Implementation**

► Results



# Learning Algorithm

## 3 Implementation



# Replay Buffer

## 3 Implementation



# Table of Contents

4 Results

- ▶ Introduction
- ▶ Algorithm comparison
- ▶ Implementation
- ▶ Results



# Hindsight Goal Prioritization for Sparse Reward Environments

*Thank you for listening!*  
*Any questions?*