ISSN print 0716-0356
ISSN online 2452-5995
ISSN-L 0716-0356

TRILOGÍA

CIENCIA · TECNOLOGÍA · SOCIEDAD

ARTÍCULOS

▪ PRESENTACIÓN

CONFERENCE PAPER

> ARTÍCULO | CONFERENCE PAPER
▪ DESIGN AND IMPLEMENTATION OF A ROBOTIC ARM CONTROLLED BY REINFORCEMENT LEARNING TECHNIQUES

> ARTÍCULO | CONFERENCE PAPER
▪ SMART IRRIGATION THROUGH WATER CONSUMPTION: PREDICTION

> ARTÍCULO | CONFERENCE PAPER
▪ STRUCTURED GRAMMATICAL EVOLUTION FOR MODELING THE MULTI-BAND LIGHT CURVES OF SUPERNOVA.

>COMUNICACIÓN BREVE
▪ JUAN JOSÉ SALDAÑA GONZÁLEZ: ADELANTADO DE LA HISTORIA DE LA CIENCIA EN AMÉRICA.

>DISCURSO
▪ ALEJANDRO *MONO* GONZÁLEZ: LA DESAPARICIÓN DE UNA EXTREMA VISIBILIDAD

UTEM
UNIVERSIDAD TECNOLÓGICA METROPOLITANA
del Estado de Chile

AGOSTO | 2023
Vol. 38 • Nº 49

EVIC 2022 UTEM
Dossier
Inteligencia Computacional

Conference paper EVIC versión XVII, 2022

Diego Quintanilla*
Universidad Tecnológica Metropolitana,
Santiago, Chile
https://orcid.org/0009-0009-6398-3490

Cristóbal Sánchez**
Universidad Tecnológica Metropolitana,
Santiago, Chile
https://orcid.org/0009-0005-6657-364X

Artículo

DISEÑO E IMPLEMENTACIÓN DE UN BRAZO ROBÓTICO CONTROLADO MEDIANTE TÉCNICAS DE APRENDIZAJE POR REFUERZO

DESIGN AND IMPLEMENTATION OF A ROBOTIC ARM CONTROLLED BY REINFORCEMENT LEARNING TECHNIQUES***

Recibido: 2 de diciembre de 2022 | Aprobado: 24 de abril de 2023 | Versión final: 16 junio de 2023

Cómo citar este artículo:
Quintanilla González, D. y Sánchez Larenas, C. (2023). Design and implementation of a robotic arm controlled by reinforcement learning techniques. Trilogía (Santiago), 38(49), 12-17. Santiago de Chile: Ediciones UTEM.

* Electronics Engineering School. Universidad Tecnológica Metropolitana, Santiago de Chile
Correo electrónico: diego.quintanillag@utem.cl
** Electronics Engineering School. Universidad Tecnológica Metropolitana, Santiago de Chile
Correo electrónico: cristobal.sanchezl@utem.cl
*** Poster Competition 2rd. XVII IEEE Latin-American Summer School on Computational Intelligence (EVIC)
December 12-14, 2022, Universidad Tecnológica Metropolitana, Chile.

ABSTRACT

The main objective of this work is the creation of a simulation from the design of a real robot controlled by RL (Reinforcement Learning), solving this through the MuJoCo physics engine, to finally be contrasted with the real robot, taking care of a key aspect in its operation: the energy consumed. In this application we will see how the manipulator works with the reward, and how we address its value per action or task performed.

Keywords: robotics, reinforcement learning, machine learning

RESUMEN

El objetivo principal de este trabajo es la creación y el diseño de un robot controlado por RL (Reinforcement Learning), resolviendo éste a través del motor de física MuJoCo, para luego ser contrastado con una implementación de un robot real en el laboratorio, cuidando un aspecto clave en su funcionamiento operativo: la energía consumida, cómo funciona y cómo abordamos su valor por acción o tarea realizada.

Palabras clave: robótica, aprendizaje por refuerzo, aprendizaje automático

I. INTRODUCTION

Nowadays, due to the most advanced development in robotic applications, it is necessary to study in a more detailedway the behavior of robotic systems and their structure, so thatit is possible to obtain predictable and increasingly sophisticated behaviors. One of the challenges is to achieve repetitive tasks in such a way that they can be implemented through software to perform various virtual simulations, optimizingthe resources available for these tasks and taking care of the integrity of people as well, because during the pandemic this became much more important, remote management and the protection of the person [1].

In this aspect, the research will aim at designing and simulating robots that have the ability to perform tasks in an autonomous and innovative way, through the integration of artificial intelligence algorithms as the main axis, taking into account and evaluating some crucial aspects, such as their operational capabilities, higher fidelity control and collision safety, etc. This will be done through repeated computational simulations, which iswhere the general study of this topic willfocus, to thenbe contrasted with an implementation of a real robot in the laboratory.

II. METHOD

A. General aspects of the framework

This environment is based on the MuJoCo physics engine [1], and for ease of implementation, we used the MuJoCo- based infrastructure called "Robosuite", containing a multi- tude of elements (robots, room, controllers, objects, etc.) that make its native structure modular, providing versatility in the creation of environments and customized robots [2].

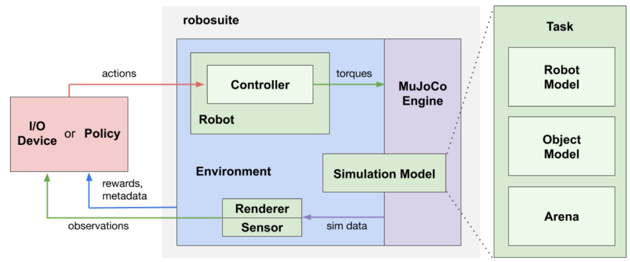


Fig. 1. Robosuite framework.

B. Task to solve

The tasks to be implemented are composed first of a robot,a table, an object on it (in this case a small cube) and atarget or reference to move towards, as well as a box on the table where the target is located. This environment can be customized with any element required, including the robot designed for this purpose. An example of this is the environments in figures 2 and 3.

In the following image the customized robot is shown in the environment also created by us, but we have not yet been able to do tests with it. In the results obtained we trained a robot called "Panda" and in others a robot exported from GYM [5] called " Fetch".

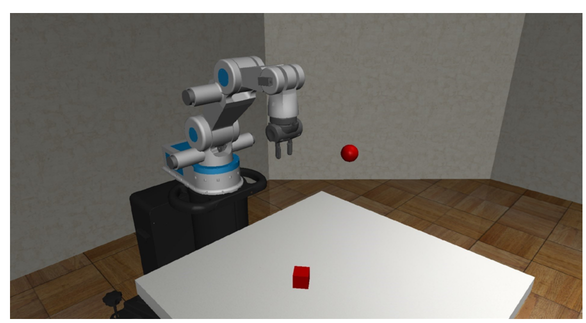


Fig. 2. Enviroment with a cube and a reference.

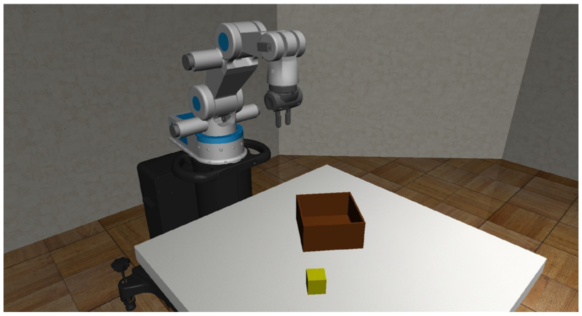


Fig. 3. Environment with a cube and a box for reference.

C. Reward and cost function

To implement the reward, several aspects were taken into account, among them is that the task is completed with a follow-up of ”steps” (stages) to which a small percentagereward is assigned to each one, which is determined accordingto the fulfillment of them; These stages are composed bythe first one, called ”Reaching” corresponds to the Euclidean distance that exists between the midpoint of the gripper andthe object (the cube on the table). Then, the next stage iscalled ”Grasping” which corresponds to when the gripper manages to grasp the object so that it can be lifted, and finally the ”Lifting” stage that would move the grasped object to the target (where there is a red ball as a reference). The following code is a brief description of how the logic of the stages is approached.

```
initialize reward to 0;
obtain the distance between the cube and
gripper mid-point;
normalize and sum to reward;
if the cube is grasped:
    sum a value to reward;
    sum the distance between the cube
    and the objective normalized to
    reward;
end if
return reward
```

D. Algorithm

For this environment, as previously mentioned, it is nec- essary to work with a continuous space algorithm, so for this case the algorithms provided by the Stable Baselines 3 library [3] were implemented, choosing DDPG, SAC and TD3,with which training was carried out to perform experiments to determine the behavior of the environment.

III. RESULTS

The following images represent different experiments with each algorithm to which the environment has been subjected, where the first one corresponds to the difference between continuous space algorithms with respect to the reward, and the second one represents the comparisons made by increasing the simulation times from 500 to 1000 timesteps, and also the differences between applying or not applying a cost functionas those presented in the figures 4 and 5.

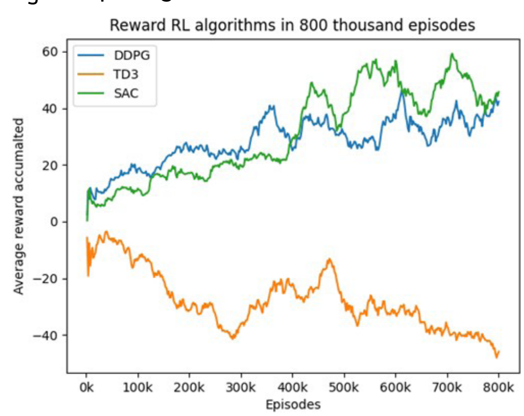


Fig. 4. Reward obtained through various algorithms.

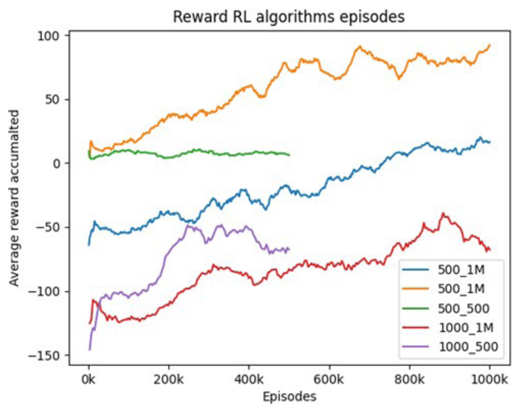


Fig. 5. Reward obtained through different episodes and timesteps.

REFERENCES

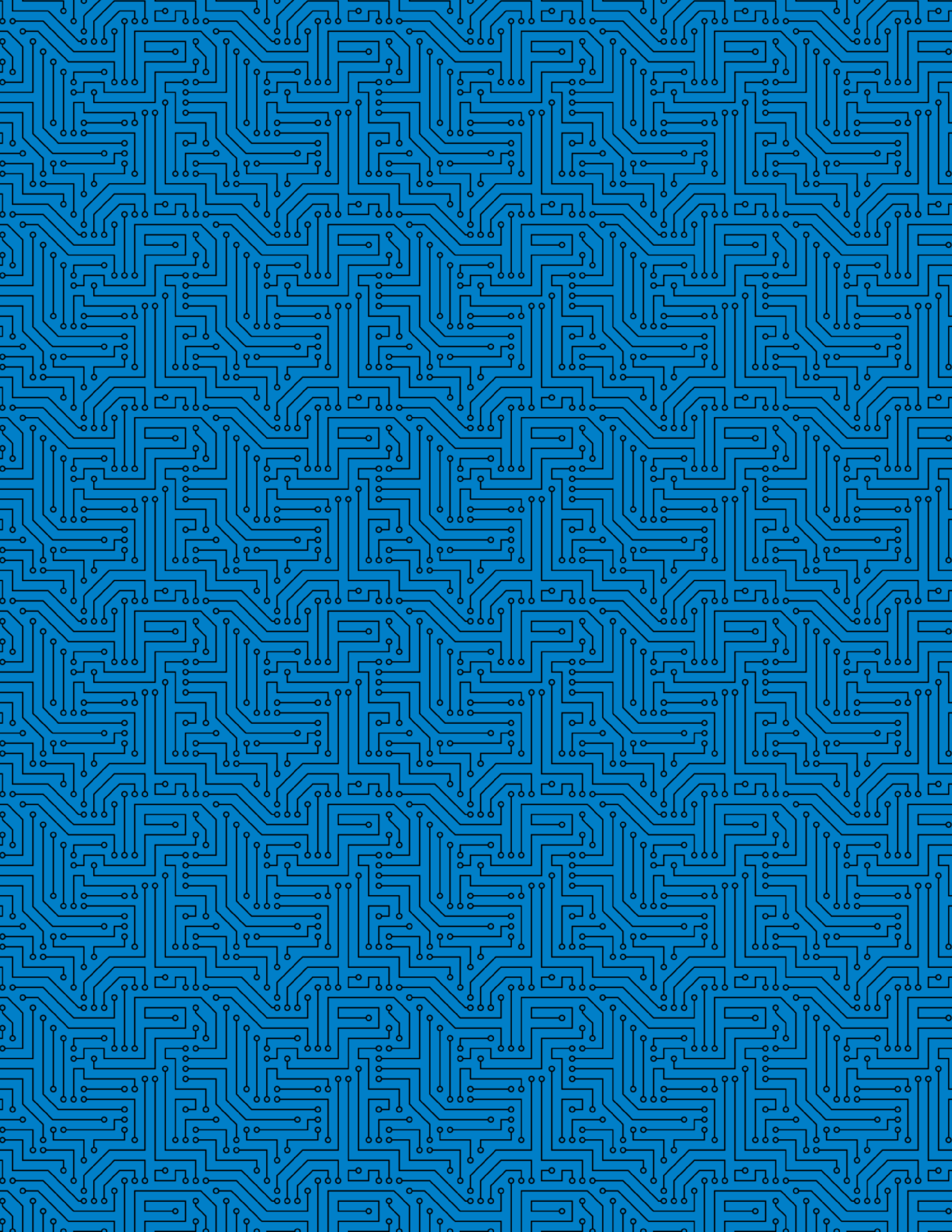
Yuan Tian, (2019). "Model Free Reinforcement Learning with Stability Guar-antee", Master of Science Thesis, Faculty of Mechanical, Maritime and Materials Engineering (3mE) Delft University of Technology, August12.

E. Todorov, T. Erez and Y. Tassa ,(2012)."MuJoCo: A physics engine for model-based control," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5026-5033, doi: 10.1109/IROS.2012.6386109.

Yuke Zhu and Josiah Wong and Ajay Mandlekar, and Abhishek Joshi and Soroush Nasiriany and Yifeng Zhuand Roberto Mart, (2009). "robosuite: A Modular Simulation Framework and Benchmark for Robot Learning", 2020, arXiv preprint arXiv:.12293.

Antonin Raffin and Ashley Hill and Adam Gleave and Anssi Kanervistoand Maximilian Ernestus and Noah Dormann, (2021). "Stable-Baselines3: Re- liable Reinforcement Learning Implementations", Journal of Machine Learning Research, vol.XX, 268, pp.1-8.

OpenAI, "Robogym" (2020). Recuperado de: https://github.com/openai/robogym.



UNIVERSIDAD TECNOLÓGICA METROPOLITANA
UTEM
del Estado de Chile

TRILOGÍA
CIENCIA · TECNOLOGÍA · SOCIEDAD

EDICIONES UNIVERSIDAD TECNOLÓGICA METROPOLITANA