

Trabajo Fin de Máster

APRENDIZAJE REFORZADO PARA AGARRE Y MANIPULACIÓN DE ROBOTS

REINFORCEMENT LEARNING FOR GRASPING AND HADLING IN ROBOTS

Autor/es

Daniel Cubel Gálvez

Director/es

Rubén Martínez Cantín

Máster en Ingeniería Industrial

Escuela de Ingeniería y Arquitectura

2020

ÍNDICE DE CONTENIDOS

[1. INTRODUCCIÓN 4](#_Toc42775459)

[1.1. OBJETO 4](#_Toc42775460)

[1.2. ALCANCE 4](#_Toc42775461)

[2. MARCO TEÓRICO 4](#_Toc42775462)

[2.1. ROBÓTICA 4](#_Toc42775463)

[2.2. APRENDIZAJE POR REFUERZO 4](#_Toc42775464)

[2.3. OPTIMIZACIÓN BAYESIANA 4](#_Toc42775465)

[3. DEFINICIÓN DEL PROBLEMA 4](#_Toc42775466)

[3.1. ENTORNO ROBÓTICO 4](#_Toc42775467)

[3.2. TAREAS 4](#_Toc42775468)

[4. SOLUCIÓN ADOPTADA 4](#_Toc42775469)

[4.1. DEFINICIÓN DE LAS TAREAS 4](#_Toc42775470)

[5. SIMULACIÓN 4](#_Toc42775471)

[6. CONCLUSIONES 4](#_Toc42775472)

# OBJETO

El objeto de este Trabajo Fin de Máster (TFM) es el desarrollo de una serie de tareas robóticas, de forma que el robot pueda aprender mediante aprendizaje por refuerzo un comportamiento deseado, diseñando la recompensa que obtiene el robot al actuar de una manera determinada al ejecutar la tarea y la política que debe seguir, que será obtenida mediante optimización bayesiana.

La justificación de este trabajo se centra sobre todo en que el problemas de las tareas robóticas en ambientes controlados, en los que se conoce la forma de los objetos que se quiere manipular, su posición y orientación, y la posición y orientación del elemento terminal está prácticamente resuelto. Por ejemplo, se puede observar en el uso de los robots en las líneas de producción. Sin embargo, al entrar en ambientes no controlados pueden aparecer problemas que antes no existían, ya que a priori no se conocen las características de los objetos que se encuentran en el entorno del robot.

El aprendizaje por refuerzo supone una revolución en muchos campos de la investigación y de la tecnología, como por ejemplo en este trabajo tareas robóticas de manipulación, pero también puede ser aplicado a una amplia variedad de problemas, como pueden ser sistemas de recomendación personalizado, diseño de estrategias financieras, juegos… en donde otros métodos fallan ya sea por la falta de estructura del entorno, la complejidad del espacio de soluciones posibles o el gran volumen de datos.

En nuestro caso, el aprendizaje por refuerzo nos va a servir para resolver una serie de tareas en las que el robot aprenderá la forma óptima de comportarse, en lugar de resolver las tareas de forma tradicional.

# ALCANCE

Las actividades desarrolladas para la realización de este trabajo fin de máster han sido:

1. Estudios previos.
   1. Revisión bibliográfica y estado del arte del aprendizaje por refuerzo.
   2. Revisión bibliográfica y estado del arte de la optimización bayesiana.
2. Familiarización con las herramientas tecnológicas.
   1. Lenguaje de programación Python.
   2. Sistema operativo Ubuntu (Linux).
   3. Plataforma de simulación de robots CoppeliaSim.
   4. API de PyRep, que funciona por encima de CoppeliaSim.
   5. Benchmark y entorno de aprendizaje RLBench, que funciona por encima de PyRep y CoppeliaSim.
3. Diseño de tareas fácilmente reproducibles para los algoritmos de aprendizaje por refuerzo.
4. Diseño de políticas para resolver las tareas planteadas.
5. Escribir el código de interfaz entre el simulador y la librería de aprendizaje.
6. Conclusiones.

# MARCO TEÓRICO

## ROBÓTICA

## APRENDIZAJE POR REFUERZO

## OPTIMIZACIÓN BAYESIANA

# DEFINICIÓN DEL PROBLEMA

## ENTORNO ROBÓTICO

## TAREAS

# SOLUCIÓN ADOPTADA

## DEFINICIÓN DE LAS TAREAS

En el simulador, en Python…

# SIMULACIÓN Y RESULTADOS

# CONCLUSIONES