



Universidad
Internacional
de Valencia

Aprendizaje por refuerzo aplicado a tareas de control

Titulación:
Máster de Inteligencia
Artificial

Curso académico
2021-2022

Alumno/a: Werner Seoane
Lucas Ezequiel
D.N.I: 39459365J

Director/a de TFM:
Gabriel Enrique Muñoz

Convocatoria:
Tercera

Mayo 2022

De:
 Planeta Formación y Universidades

*Escribe aquí
tu frase favorita.*

E indica aquí su autor

Agradecimientos

A mi familia. A mis padres y mis hermanos y hermanas.

Índice general

Índice de figuras	III
Índice de tablas	IV
Índice de algoritmos	V
Resumen	1
1. Introducción	3
1.1. Acotación del problema	3
1.2. Dispositivo utilizado	3
1.3. Marco teórico	3
1.3.1. Aprendizaje por refuerzo	3
1.3.2. Algoritmo Policy Gradient	3
1.3.3. Algoritmo Actor Critic	3
1.3.4. Vision Transformers	3
1.3.5. Una subsección	4
1.3.6. Una subsubsección	4
2. Objetivos	7
3. Metodología	9
3.1. Obtención de los datos	9
3.2. Preprocesamiento de los datos	9
3.3. Análisis e investigación de soluciones de aprendizaje por refuerzo	9
3.3.1. Elección de los algoritmos utilizados por el agente	9
3.3.2. Definición del agente	9
3.3.3. Definición del estado	9
3.3.4. Definición del estado	9
3.3.5. Definición del entorno	9
3.3.6. Entrenamiento	9
3.3.7. Evaluación	9

3.4.	9
4. Resultados y Discusión	11
5. Conclusiones	12
6. Limitaciones y Perspectivas de Futuro	13
A. Apéndice A	16
B. Apéndice B	17
Bibliografía	18



Índice de figuras

1.1. Tipos de grafos 4



Índice de tablas

1.1. Ejemplo de tabla 4

Índice de algoritmos

1.	Algoritmo <i>Hill-Climbing</i> (HC)	5
----	---	---



Resumen

Introducción

1

Escribe aquí la introducción de tu Trabajo Fin de Máster, utilizando tantas secciones, subsecciones y subsubsecciones como estimes necesarias.

1.1. Acotación del problema

1.2. Dispositivo utilizado

Esta palabra está en negrita. *Esta palabra* está en cursiva. **Esta palabra** se destaca en púrpura.

1.3. Marco teórico

1.3.1. Aprendizaje por refuerzo

1.3.2. Algoritmo Policy Gradient

1.3.3. Algoritmo Actor Critic

1.3.4. Vision Transformers

En la sección ?? se muestran ejemplos de palabras en negrita, cursiva y destacadas en púrpura.

Una Red Generativa Antagónica o *Generative Adversarial Network* (GAN) es... ([Goodfellow et al., 2014](#)).

[Goodfellow et al. \(2014\)](#) diseñaron las redes generativas antagónicas como...

Listado:

- Item 1.
- Item 2.
- Item 3.

Enumeración:

1. Item 1.

2. Item 2.

3. Item 3.

1.3.5. Una subsección

La figura 1.1 muestra...

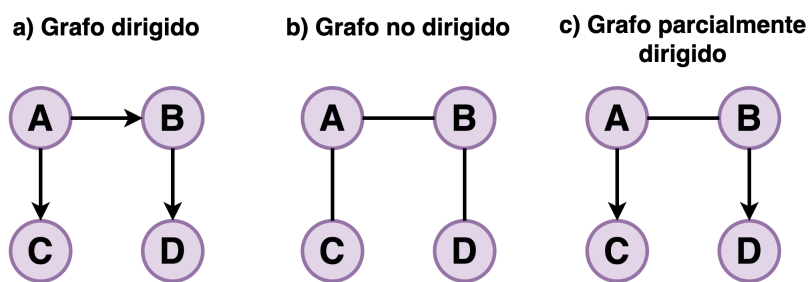


Figura 1.1: *Tipos de grafos.*

La tabla 1.1 muestra...

Columna 1	Columna 2	Columna 3	Columna 4	Columna 5
Fila 1	A	B	C	D
Fila 2	E	F	G	H
Fila 3	I	J	K	L

Tabla 1.1: *Ejemplo de tabla.*

1.3.6. Una subsubsección

El algoritmo 1 muestra...

Algoritmo 1: Algoritmo *Hill-Climbing* (HC)

1. Elegir una estructura de red \mathcal{G} sobre \mathbf{V} , normalmente vacía. Establecer la puntuación máxima inicial: $Score_{max} = Score_{\mathcal{G}}$.
 2. Repetir los siguientes pasos mientras $Score_{max}$ siga aumentando:
 - a) Calcular las puntuaciones para todas las posibles redes modificadas \mathcal{G}^* que se pueden obtener añadiendo, eliminando o reorientando un solo eje de \mathcal{G} sin que se produzcan ciclos.
 - b) Si para alguna de las redes modificadas \mathcal{G}^* se cumple que $Score_{\mathcal{G}^*} > Score_{\mathcal{G}}$, establecer $G = G^*$ y $Score_{max} = Score_{\mathcal{G}^*}$.
 3. Devolver el DAG \mathcal{G} .
-

Ejemplo de fórmula:

$$N_k(\mu, \Sigma) = \frac{1}{\sqrt{2\pi \det(\Sigma)}} \exp \left\{ -\frac{1}{2}(\mathbf{X} - \mu)^T \Sigma^{-1}(\mathbf{X} - \mu) \right\} \quad \mathbf{X}, \mu \in \mathbb{R}^k$$

Otro ejemplo de fórmula:

$$\underbrace{P(\mathcal{B}|\mathcal{D}) = P(\mathcal{G}, \Theta|\mathcal{D})}_{\text{Aprendizaje}} = \underbrace{P(\mathcal{G}|\mathcal{D})}_{\text{Aprendizaje estructural}} \cdot \underbrace{P(\Theta|\mathcal{G}, \mathcal{D})}_{\text{Aprendizaje paramétrico}}$$

Objetivos

2

Describe aquí el objetivo general de tu Trabajo Fin de Máster y, a continuación, define los objetivos parciales:

1. **Objetivo parcial 1.**
2. **Objetivo parcial 2.**
3. **Objetivo parcial 3.**

Metodología

3

3.1. Obtención de los datos

3.2. Preprocesamiento de los datos

3.3. Análisis e investigación de soluciones de aprendizaje por refuerzo

3.3.1. Elección de los algoritmos utilizados por el agente

3.3.2. Definición del agente

3.3.3. Definición del estado

3.3.4. Definición del estado

3.3.5. Definición del entorno

3.3.6. Entrenamiento

3.3.7. Evaluación

3.4.

Resultados y Discusión

4

Conclusiones

5

1. Conclusión 1.
2. Conclusión 2.
3. Conclusión 3.

Limitaciones y Perspectivas de Futuro

6

Apéndice A



Apéndice B

B

Bibliografía

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., y Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.