



UNIVERSITAT OBERTA DE CATALUNYA (UOC)
MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS (*DATA SCIENCE*)

TRABAJO FINAL DE MÁSTER

ÁREA: REINFORCED LEARNING

Impacto de la complejidad de las observaciones en el rendimiento de algoritmos de Aprendizaje por Refuerzo en un entorno Pacman

Autor: Alejandro Suau Ruiz

Tutor: Marc Borrás Camarasa

Profesor: David Masip Rodó

Palma de Mallorca, 6 de octubre de 2025

Créditos/Copyright

Una página con la especificación de créditos/copyright para el proyecto (ya sea aplicación por un lado y documentación por el otro, o unificadamente), así como la del uso de marcas, productos o servicios de terceros (incluidos códigos fuente). Si una persona diferente al autor colaboró en el proyecto, tiene que quedar explicitada su identidad y qué hizo.

A continuación se ejemplifica el caso más habitual, aunque se puede modificar por cualquier otra alternativa:



Esta obra está sujeta a una licencia de Reconocimiento - NoComercial - SinObraDerivada
3.0 España de Creative Commons.

FICHA DEL TRABAJO FINAL

Título del trabajo:	Impacto de la complejidad de las observaciones en el rendimiento
Nombre del autor:	Alejandro Suau Ruiz
Nombre del colaborador/a docente:	Marc Borrás Camarasa
Nombre del PRA:	David Masip Rodó
Fecha de entrega (mm/aaaa):	01/2026
Titulación o programa:	Máster Universitario de Data Science
Área del Trabajo Final:	Reinforced Learning
Idioma del trabajo:	Español
Palabras clave	Aprendizaje por refuerzo, Complejidad del estado, Pacman

Dedicatoria/Cita

Breves palabras de dedicatoria y/o una cita.

Agradecimientos

Si se considera oportuno, mencionar a las personas, empresas o instituciones que hayan contribuido en la realización de este proyecto.

Abstract

This work investigates how the complexity of state representation affects the performance of reinforcement learning algorithms in a simplified Pacman environment. The central idea is that an agent’s success depends not only on the algorithm itself but also on the quality and richness of the information it receives from the environment. To analyze this, we designed a set of progressively more complex observation spaces, ranging from the basic positions of the player and the ghost to configurations that include the presence and duration of power-ups and the distribution of coins across quadrants. Experiments were carried out using widely adopted algorithms such as PPO, A2C, and DQN, implemented with the Stable-Baselines3 library. Controlled training with different random seeds and consistent performance metrics allowed us to compare both learning speed and the stability of the learned policies. The results show that increasing state complexity does not always lead to better performance: in some cases, the additional information introduces noise and hinders convergence. However, for specific configurations, the agent was able to exploit the extra information to achieve more robust behavior. In conclusion, the project highlights the importance of observation design in the success of reinforcement learning agents, stressing the need to balance simplicity and expressiveness depending on the task and the chosen algorithm.

Keywords: Reinforcement learning, State complexity, Pacman, Stable-Baselines3, Gymnasium

Resumen

Este trabajo explora cómo la complejidad de la representación del estado influye en el rendimiento de los algoritmos de aprendizaje por refuerzo en un entorno simplificado de Pacman. Partimos de la premisa de que un agente no aprende únicamente por el algoritmo empleado, sino también por la calidad y la riqueza de la información que percibe del entorno. Para analizarlo, hemos diseñado un conjunto de observaciones progresivamente más complejas, desde la posición básica de jugador y fantasma, hasta configuraciones que incluyen la presencia y duración de comodines o la distribución de monedas por cuadrantes. La experimentación se ha llevado a cabo con algoritmos ampliamente utilizados, como PPO, A2C y DQN, implementados con la librería Stable-Baselines3. Se han realizado entrenamientos controlados con semillas distintas y métricas de rendimiento homogéneas, lo que ha permitido comparar tanto la velocidad de aprendizaje como la estabilidad de las políticas aprendidas. Los resultados muestran que un aumento en la complejidad del estado no garantiza siempre un mejor desempeño: en ciertos casos, la mayor riqueza de información introduce ruido y dificulta la convergencia. Sin embargo, para determinadas configuraciones el agente logra aprovechar la información adicional y alcanzar un comportamiento más robusto. En conclusión, el proyecto confirma la relevancia del diseño de observaciones en el éxito de un agente de aprendizaje por refuerzo, subrayando la necesidad de equilibrar simplicidad y expresividad en función del objetivo y del algoritmo empleado.

Palabras clave: Aprendizaje por refuerzo, Complejidad del estado, Pacman, Stable-Baselines3, Gymnasium

Índice general

Abstract	ix
Resumen	xi
Índice	xiii
Lista de Figuras	xv
Lista de Tablas	1
1. Introducción	1
1.1. Contexto y motivación	1
1.2. Objetivos	1
1.3. Sostenibilidad, diversidad y desafíos ético/sociales	2
1.4. Enfoque y metodología	2
1.5. Planificación	2
1.6. Resumen de los productos del proyecto	3
1.7. Breve descripción de los demás capítulos del informe	3
2. Métodos y recursos	3
3. Resultados	3
4. Conclusiones y trabajo futuro	3
5. Glosario	4
Bibliografía	4
6. Anexos	5

Índice de figuras

1. Ejemplo de figura. Será indexada en la “Lista de Figuras”. 1

Índice de cuadros

1. Introducción

Esta plantilla pretende ser una guía para los estudiantes. Esta plantilla se puede adaptar a las necesidades específicas de cada proyecto si el supervisor del proyecto está de acuerdo con los cambios.



Figura 1: Ejemplo de figura. Será indexada en la “Lista de Figuras”.

1.1. Contexto y motivación

Punto de partida del proyecto (¿Cuál es el problema que necesita ser resuelto? ¿Por qué es un tema relevante? ¿Cómo se está resolviendo el problema actualmente?) y descripción de la contribución (¿Qué resultado se espera obtener?)

1.2. Objetivos

Listado de los objetivos a alcanzar en este proyecto.

1.3. Sostenibilidad, diversidad y desafíos ético/sociales

Esta sección debe evaluar el impacto positivo/negativo del proyecto en las siguientes dimensiones. No es necesario alcanzar un impacto positivo en todas las dimensiones, pero es necesario considerar y discutir si existe un impacto desde el inicio del proyecto.

Sostenibilidad En el desarrollo del proyecto o durante todo su ciclo de vida (por ejemplo, despliegue, retiro), ¿tiene el resultado de este proyecto un impacto en la sostenibilidad y/o huella ecológica (consumo/ahorro de energía/recursos, desperdicio, contaminación, agotamiento de materias primas)? ¿Está afectado por leyes o regulaciones sobre este asunto? Considerando otra perspectiva, ¿afecta a alguno de los Objetivos de Desarrollo Sostenible (ODS) relacionados con estas dimensiones? Si no tiene ningún impacto, ya sea positivo o negativo, debe explicar cómo llegó a esta conclusión y justificar su respuesta.

Comportamiento ético y responsabilidad social ¿Es el resultado del proyecto demasiado técnico para tener algún impacto positivo/negativo en aspectos éticos/sociales? ¿Tiene un impacto en leyes/regulaciones (datos, privacidad, trabajo, propiedad intelectual, seguridad personal, ...)? ¿Se adhiere a los principios deontológicos de la profesión? ¿Pone en peligro/mejora/empeora algún puesto de trabajo? Si no tiene ningún impacto, ya sea positivo o negativo, debe explicar cómo llegó a esta conclusión y justificar su respuesta.

Diversidad, género y derechos humanos ¿Es el resultado de este proyecto tan técnico que no tiene impacto positivo/negativo en términos de género, diversidad o derechos humanos? ¿Y en leyes/regulaciones? ¿Y en términos de accesibilidad, discapacidad, ergonomía y/o seguridad de la información? Si no tiene ningún impacto, ya sea positivo o negativo, debe explicar cómo llegó a esta conclusión y justificar su respuesta.

1.4. Enfoque y metodología

Describa las estrategias potenciales para desarrollar este proyecto y explique la estrategia seleccionada. Discuta por qué la estrategia seleccionada es la más adecuada para alcanzar los objetivos del proyecto.

1.5. Planificación

Descripción de los recursos necesarios para desarrollar el proyecto, las tareas a realizar y una planificación temporal de cada tarea utilizando un diagrama de Gantt o equivalente. Esta planificación debe definir los hitos que se completarán en cada Prueba de Evaluación Continua (PEC).

1.6. Resumen de los productos del proyecto

No es necesario describir cada producto en detalle: esto se hará en los capítulos restantes del proyecto.

1.7. Breve descripción de los demás capítulos del informe

Breve descripción de los contenidos de cada capítulo y su relación con el resto del proyecto.

2. Métodos y recursos

En estas secciones, es necesario describir:

- Los aspectos más relevantes del diseño y desarrollo del proyecto.
- La metodología utilizada en el proceso de desarrollo, describiendo las alternativas posibles, las decisiones que se han tomado y los criterios utilizados para tomar estas decisiones.
- Una descripción de los productos que se han creado.

La estructura de estas secciones puede cambiar según el tipo de proyecto que se esté desarrollando.

3. Resultados

Describe los resultados obtenidos utilizando la metodología descrita anteriormente.

4. Conclusiones y trabajo futuro

Esta sección debe incluir lo siguiente:

- Una descripción de las conclusiones del trabajo.
- Una evaluación crítica del grado de logro de los objetivos iniciales.
- Una evaluación crítica de la planificación y metodología utilizadas en el proyecto.
- Considerando los desafíos de sostenibilidad, diversidad y ético-sociales vinculados al proyecto.
- Una discusión de temas para trabajo futuro potencial que no se hayan explorado en este proyecto.

5. Glosario

Definición de los términos y acrónimos más relevantes utilizados en este informe.

Bibliografía

Latanya Sweeney. k-anonymity: a model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10:557–570, October 2002. ISSN 0218-4885. doi: 10.1142/S0218488502001648. URL <http://portal.acm.org/citation.cfm?id=774544.774552>.

6. Anexos

Lista de secciones que son demasiado largas para ser incluidas en el cuerpo del informe y que son autocontenidas.