

Aprendizaje por Refuerzo Profundo Multiagente



Universidad
Internacional
de Valencia

Pedro Luis Soto Santos

De:



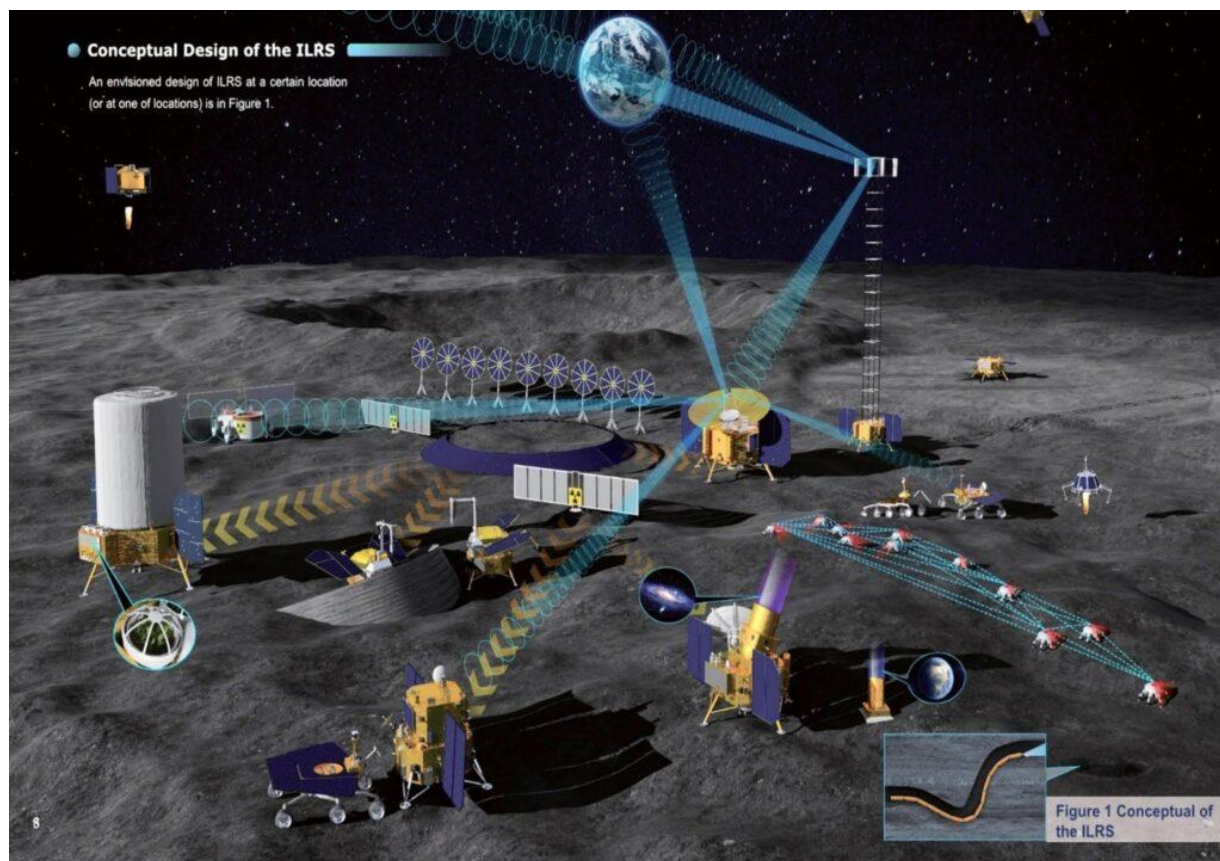
Planeta Formación y Universidades

Índice de la Presentación

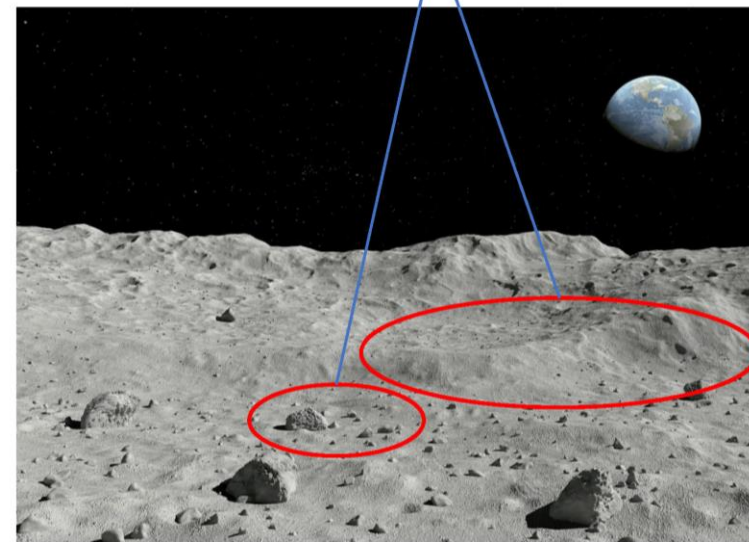
- 1 Introducción
- 2 Objetivos
- 3 Estado del Arte
- 4 Marco Teórico
- 5 Caso de Estudio
- 6 Experimentación y Pruebas
- 7 Conclusiones y Trabajo Futuro

Introducción

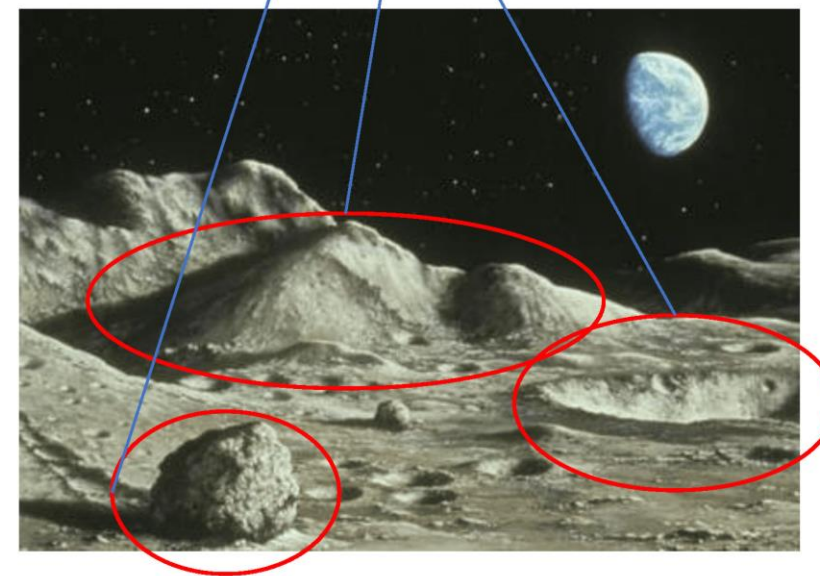
Introducción



Obstáculos pequeños



Obstáculos grandes



Problemas Multiagente

Tipos de problemas:

- Cooperativos
- Competitivos
- Mixtos

Secuencia de acciones:

- Simultanea
- Cíclica o Por Turnos
- Asíncrona

Toma de Decisiones:

- Centralizada
- Descentralizada
- Híbrida

Objetivos

Desarrollar una solución de planificación de rutas multiagente basada en algoritmos de Aprendizaje por Refuerzo Profundo

- Revisar el estado del arte sobre los algoritmos
- Desarrollar los algoritmos más destacados
- Diseñar y desarrollar un simulador
- Entrenar los algoritmos desarrollados con el simulador
- Evaluar el modelo resultante sobre entornos de prueba

Estado del Arte

Método Analíticos o Algoritmos de Búsqueda Deterministas

- A*
- Optimal Reciprocal Collision Avoidance (ORCA)

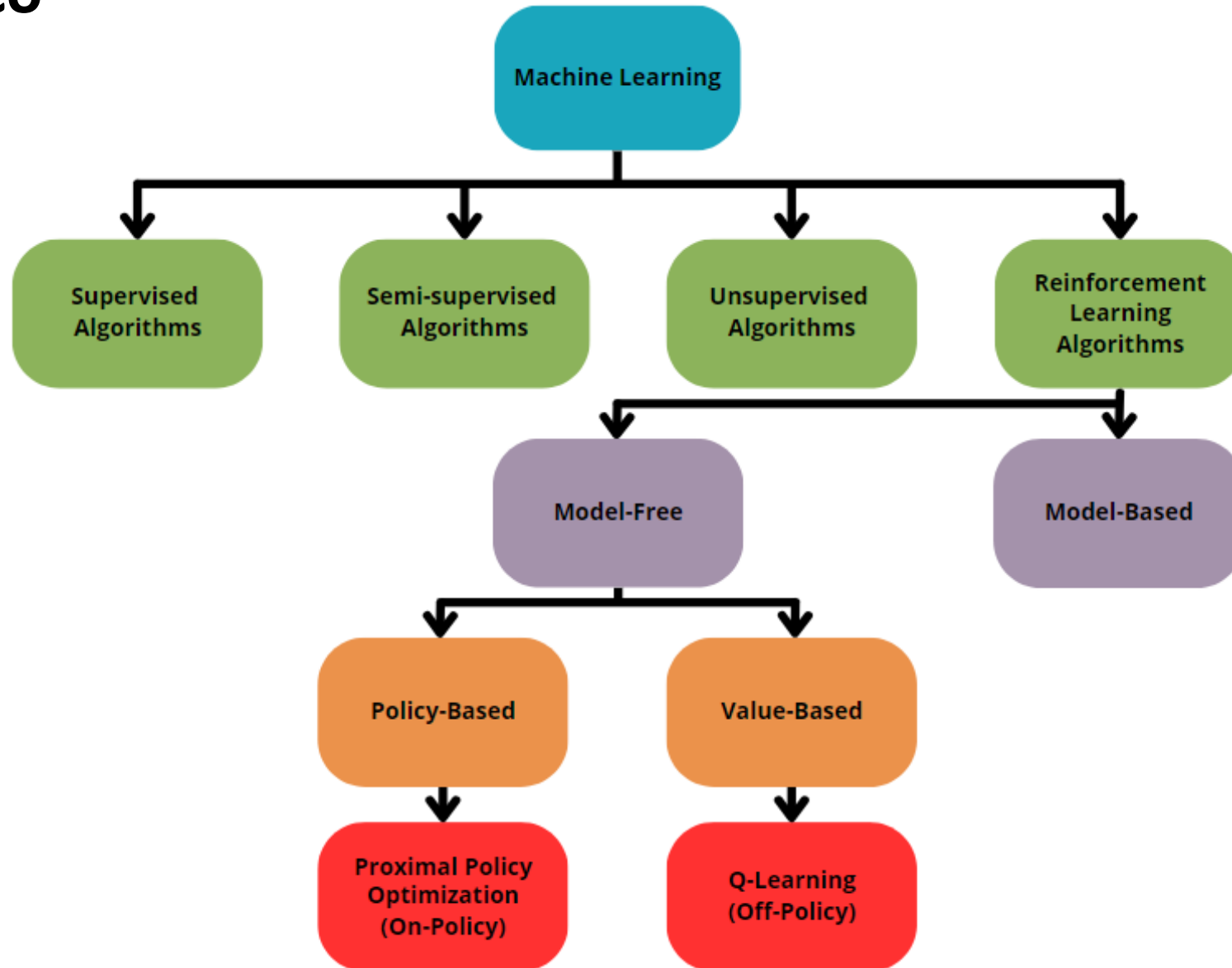


Aprendizaje por Refuerzo Profundo

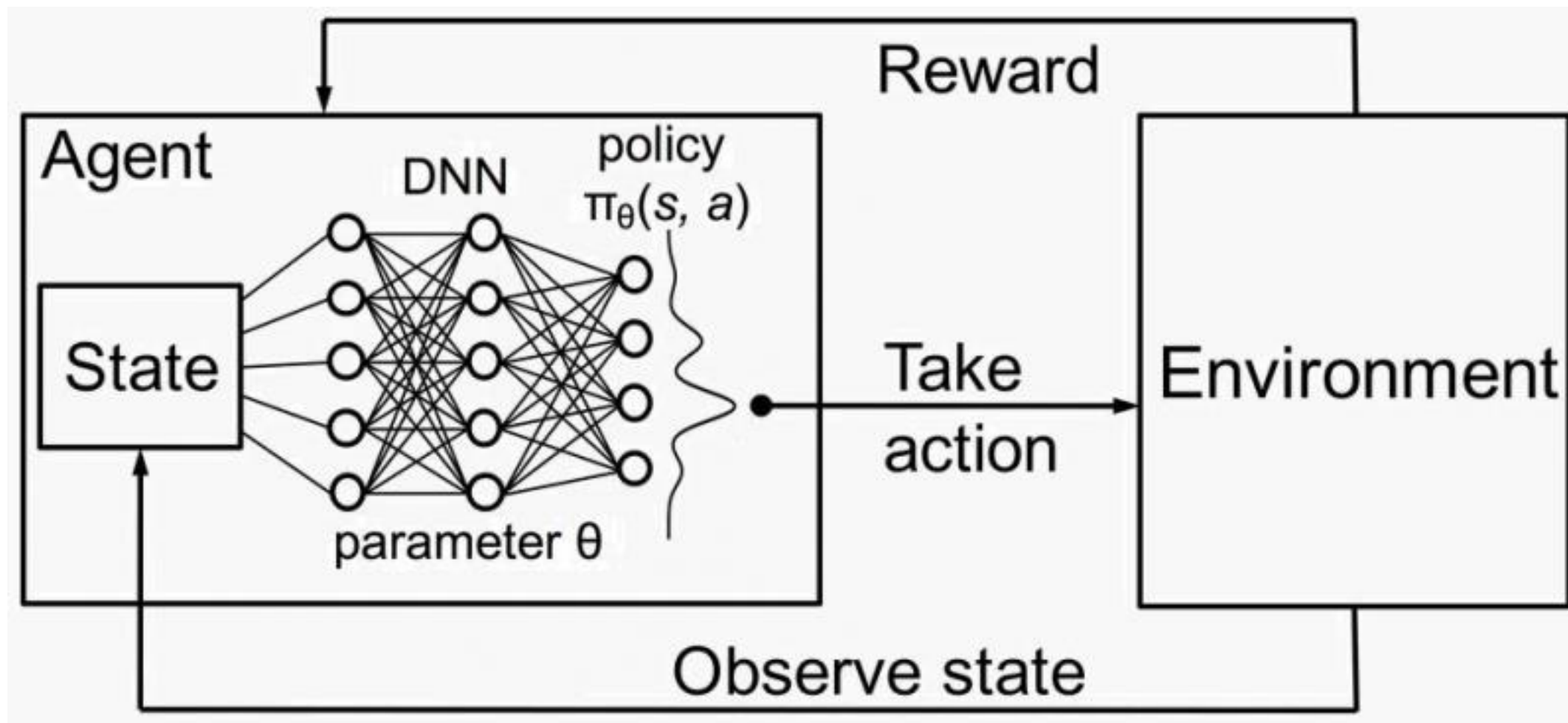
- Double Dueling Deep Q-Learning (DDDQL)
- Rainbow Deep Q-Learning (Rainbow)
- Multi-Agent Proximal Policy Optimization (MAPPO)
- Independent Proximal Policy Optimization (IPPO)

Marco Teórico

Marco Teórico

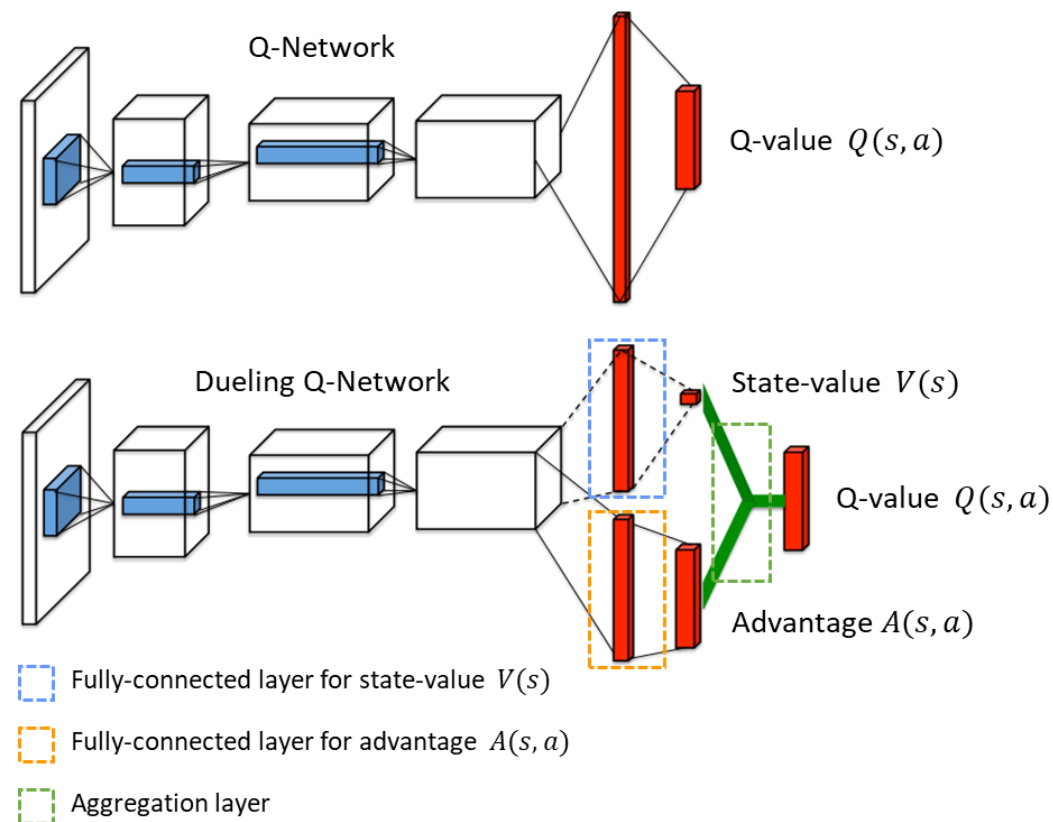
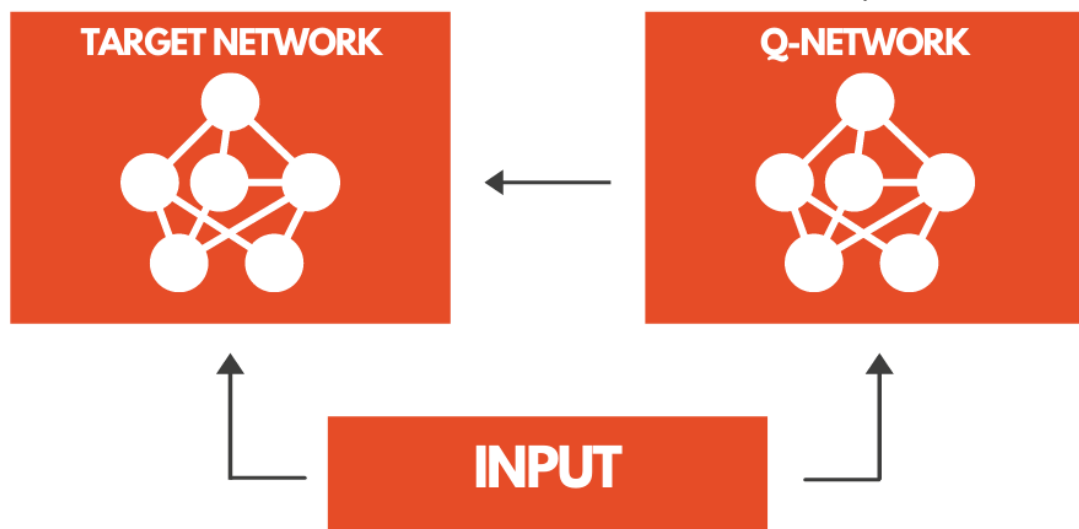


Aprendizaje por Refuerzo Profundo

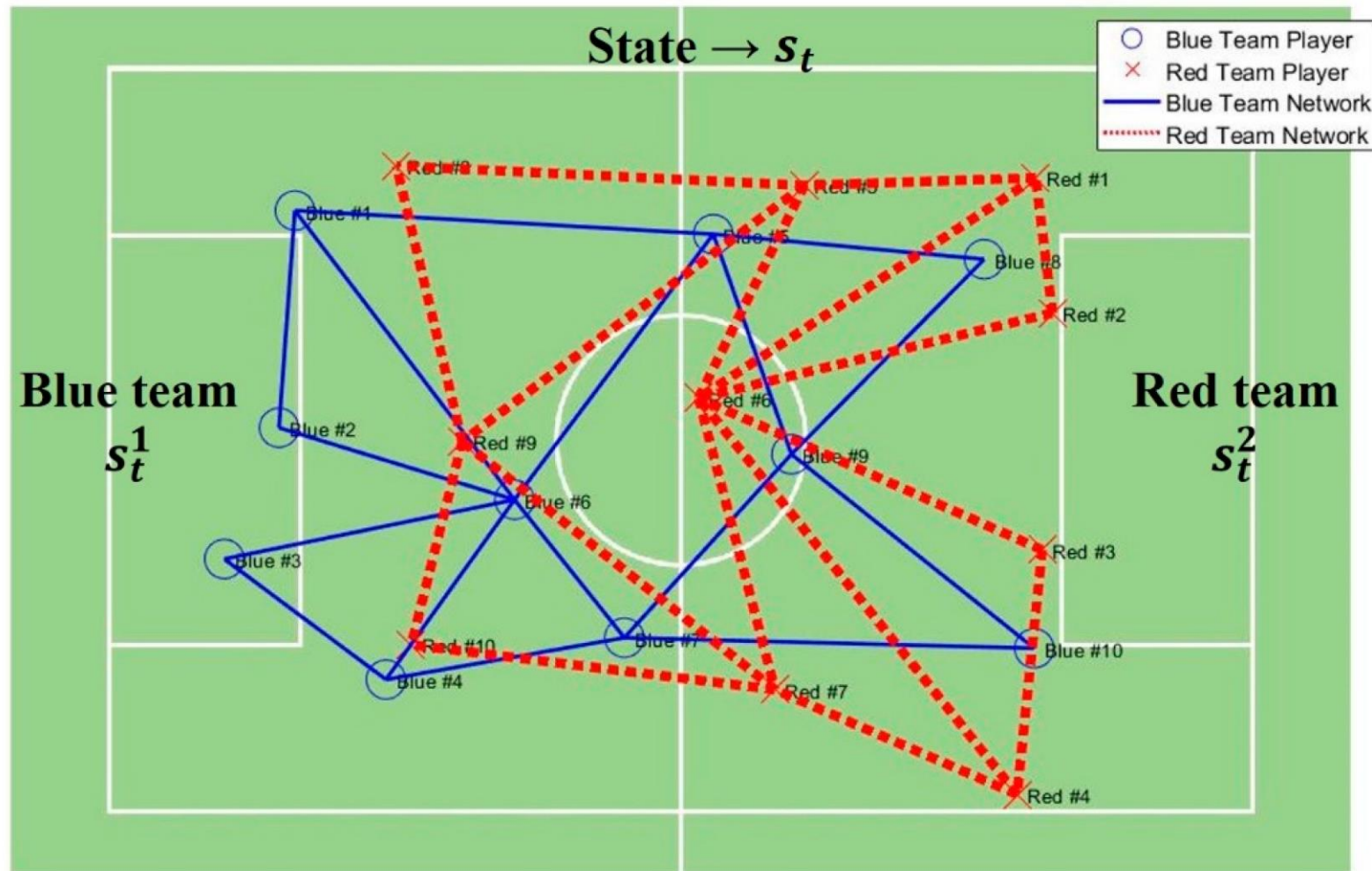


Double Dueling Deep Q-Learning

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t) \right]$$

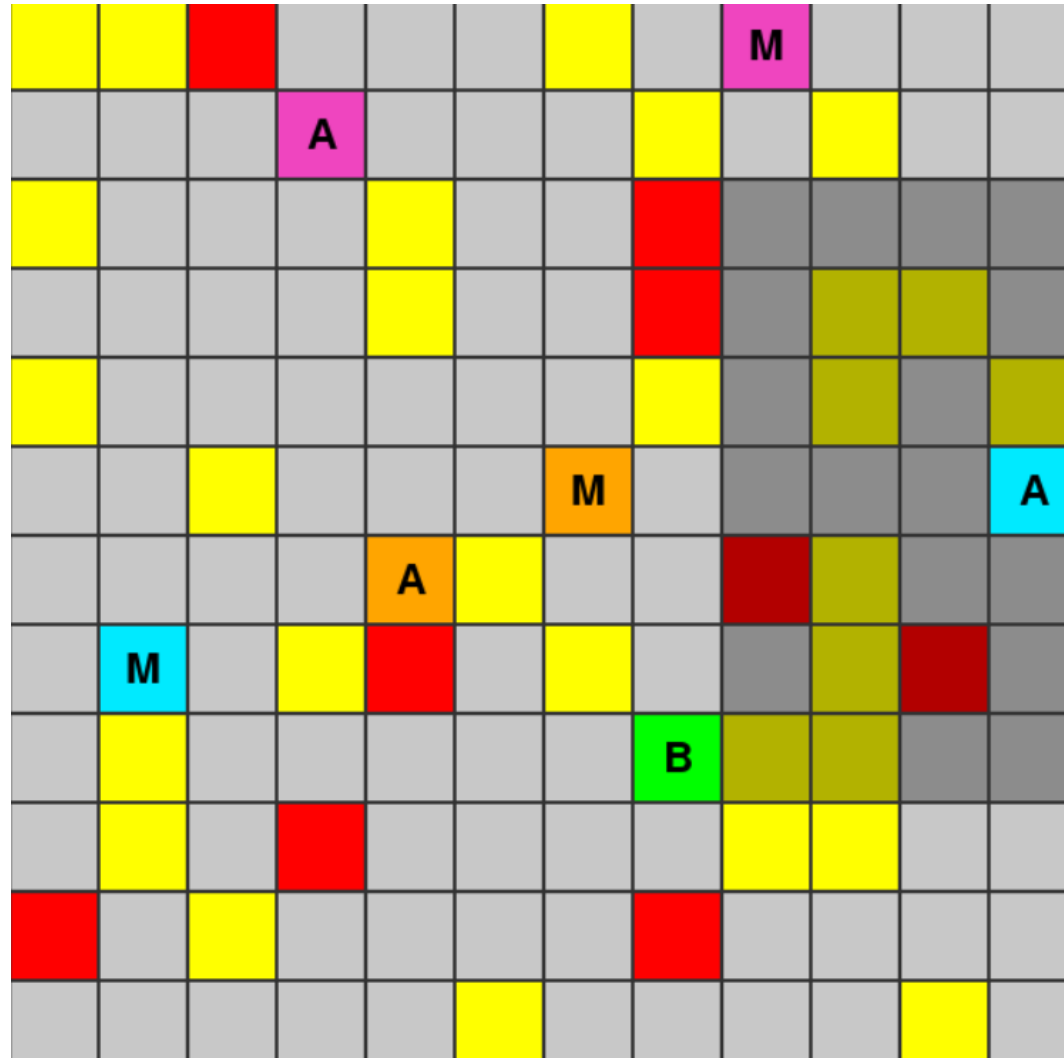


Multi-Agent Proximal Policy Optimization



Caso de Estudio

Simulación



Casilla **Amarilla** = Obstáculo pequeño

Casilla **Roja** = Obstáculo grande

Casilla **Gris** = Espacio vacío

Casilla **Verde** ("B") = Mezcladora

Casilla **Azul** ("A") = Agente 1

Casilla **Azul** ("M") = Mina 1

Casilla **Oscura** = Rango de visión del agente 1

Casilla **Naranja** ("A") = Agente 2

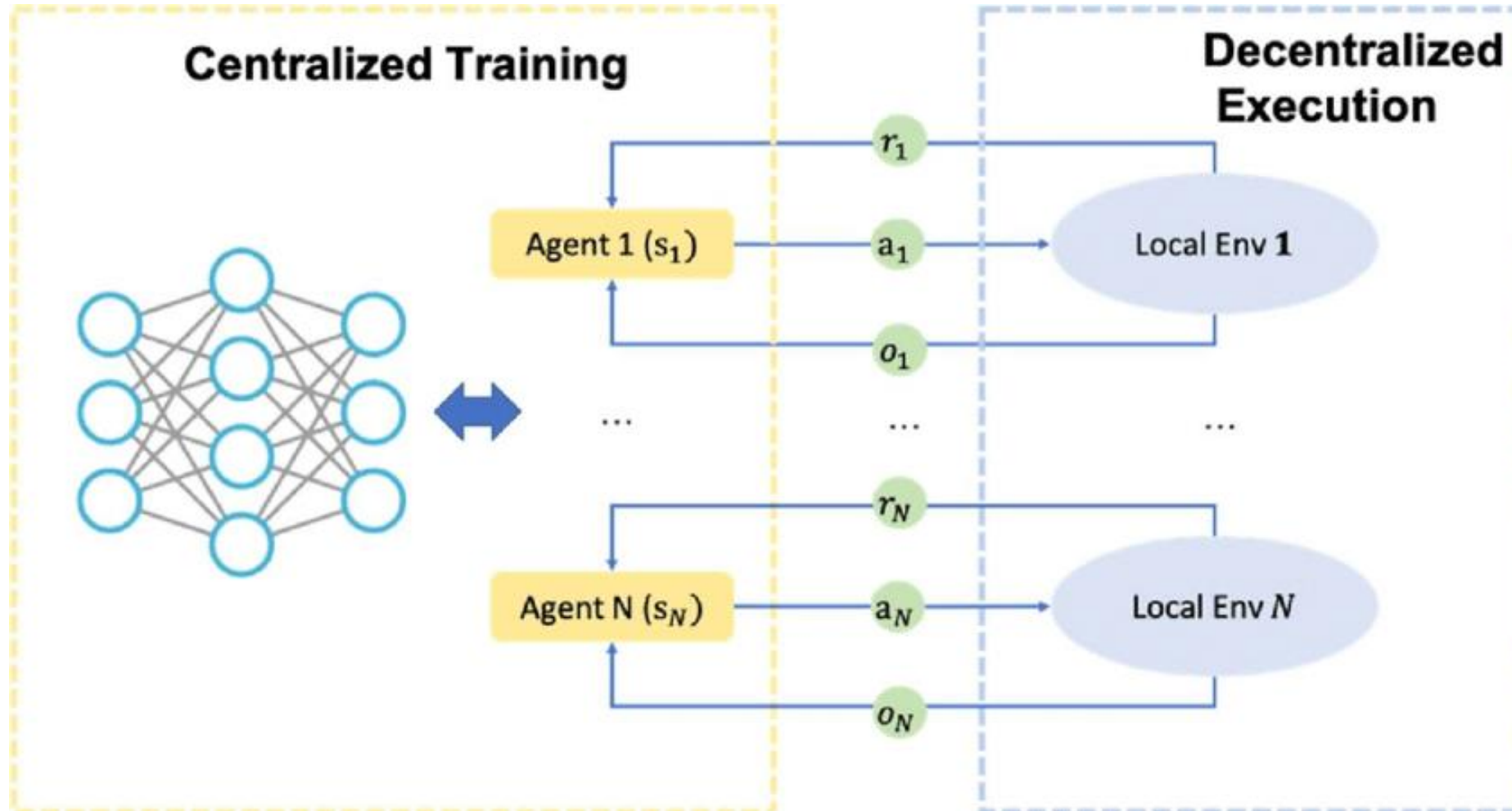
Casilla **Naranja** ("M") = Mina 2

Casilla **Rosa** ("A") = Agente 3

Casilla **Rosa** ("M") = Mina 3

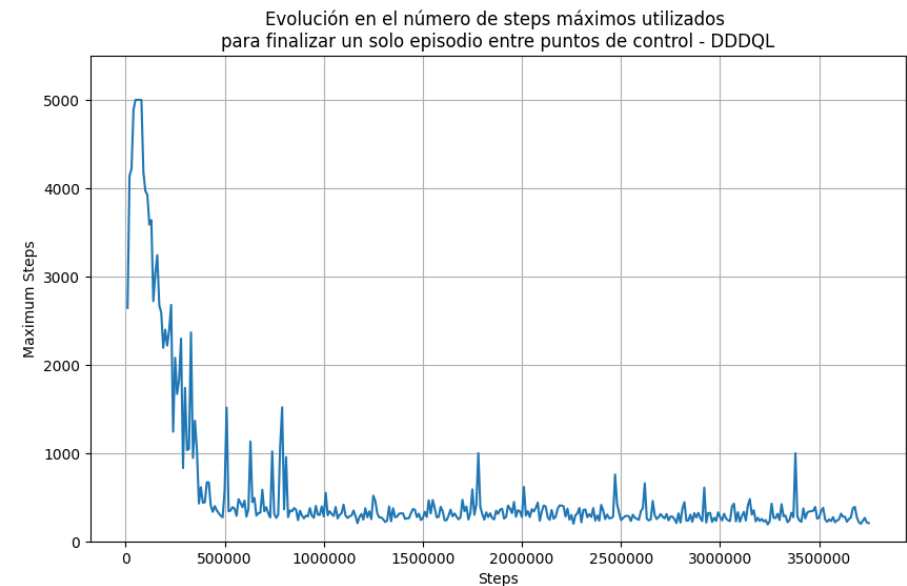
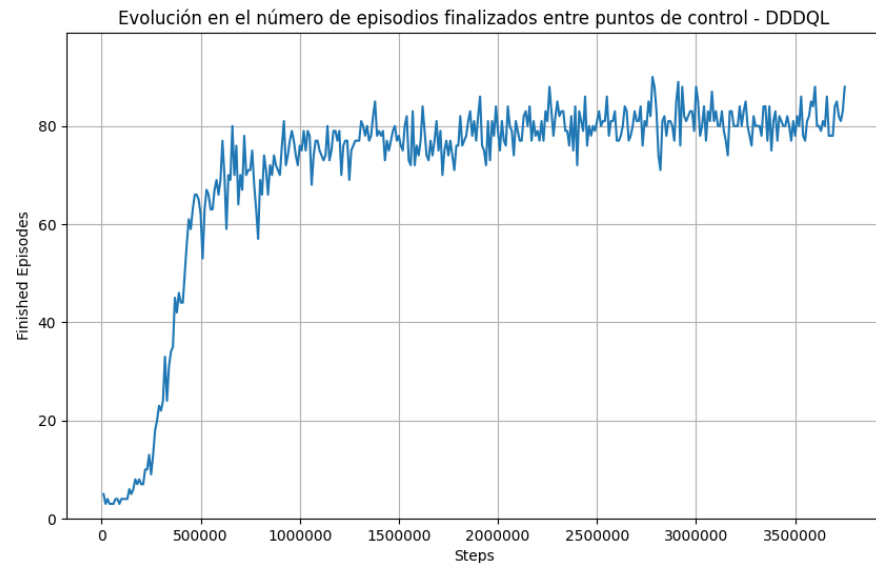
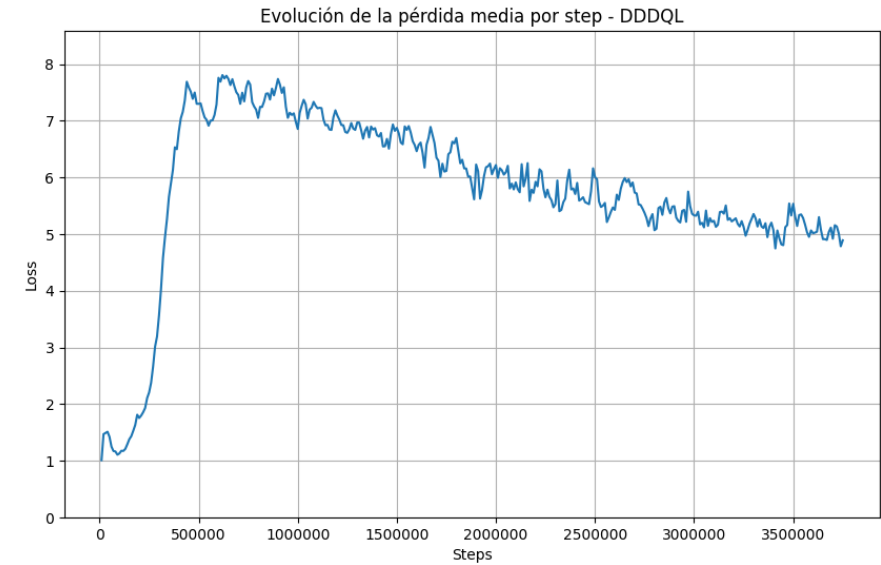
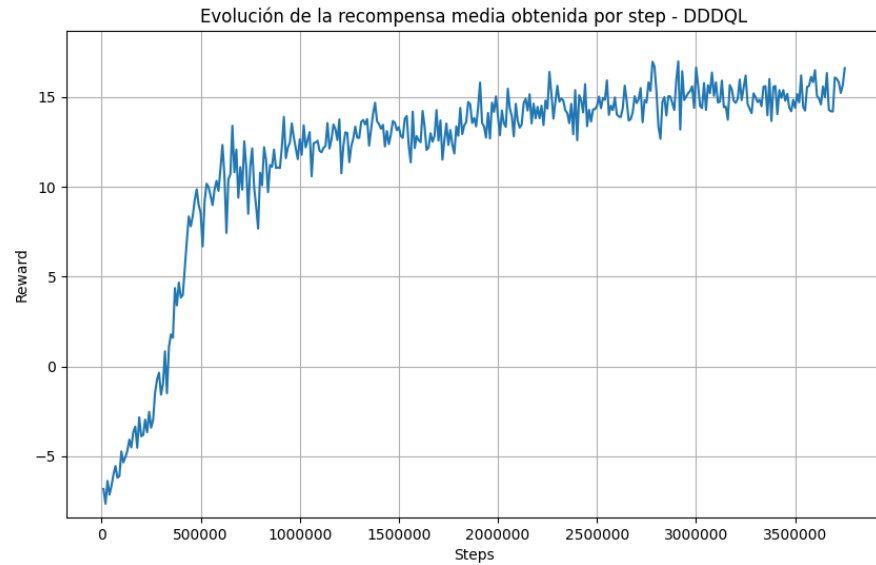
Agentes Con Conocimiento Previo VS Agentes Sin Conocimiento Previo

Caso de Estudio

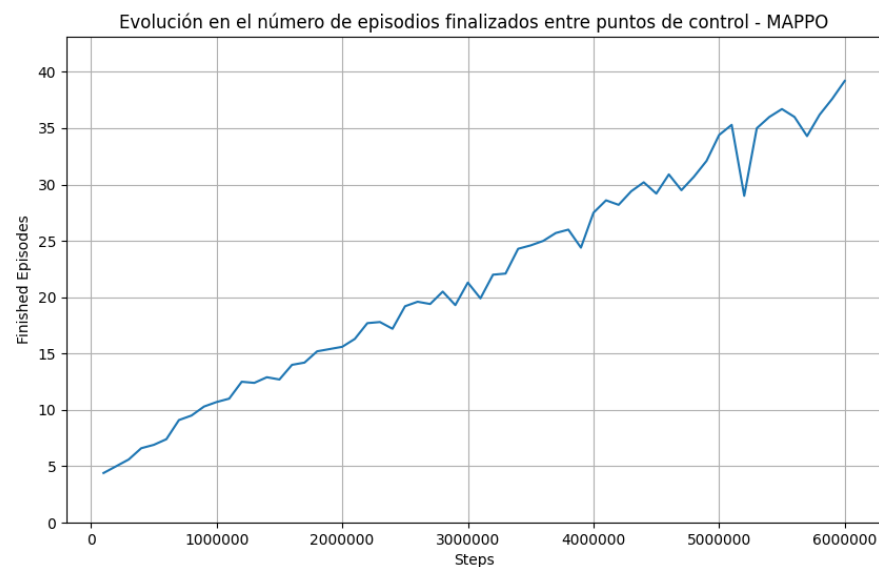
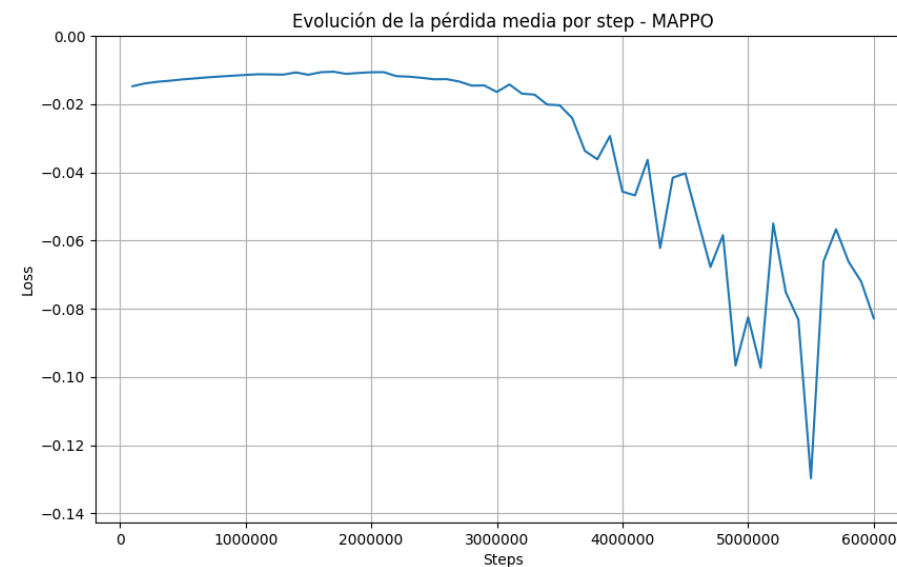


Experimentación y Pruebas

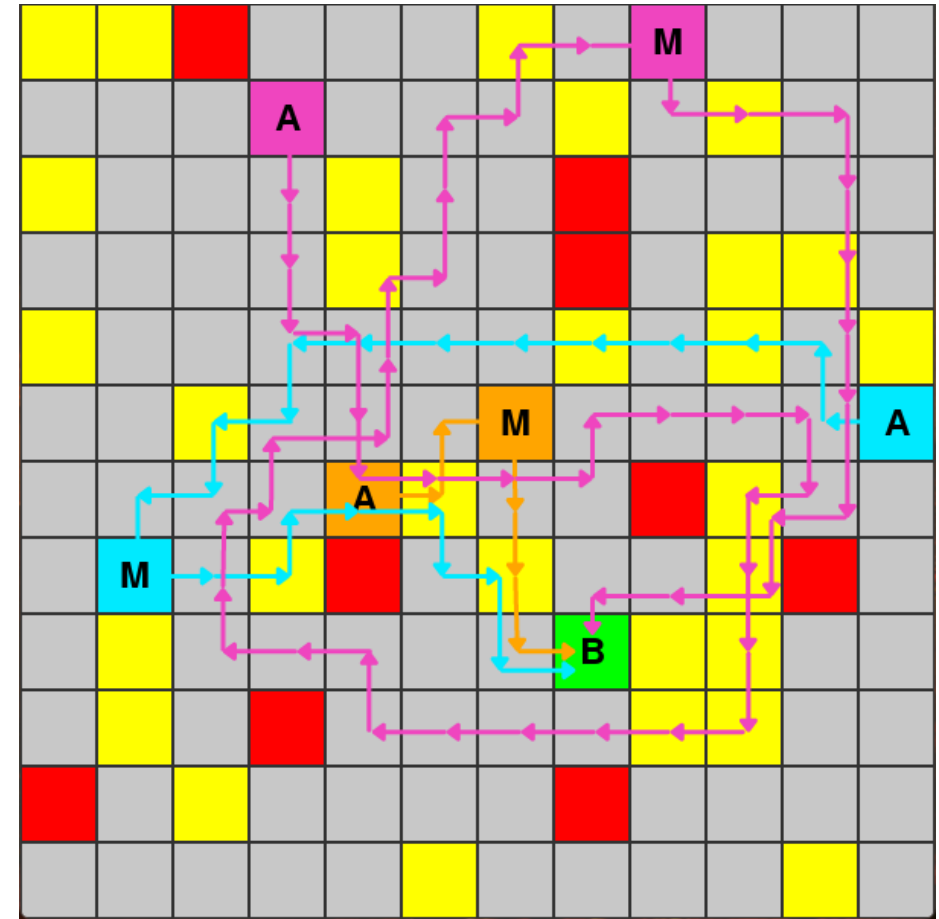
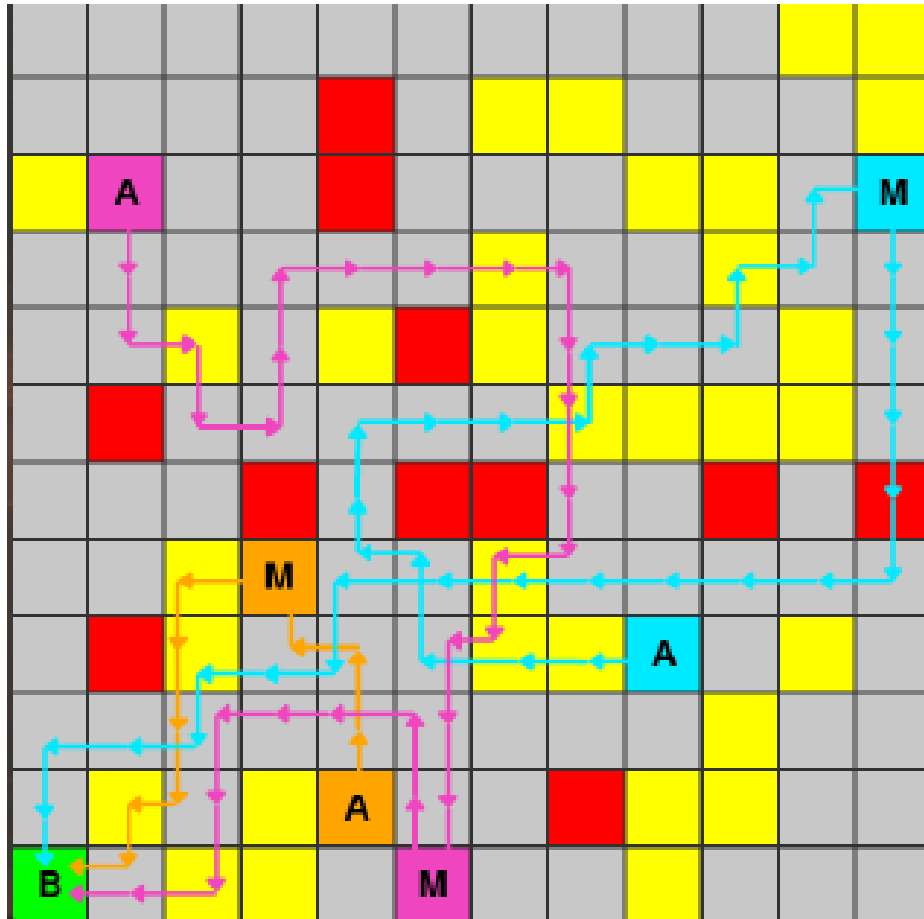
Experimentación y Pruebas Evolución Entrenamiento DDDQL



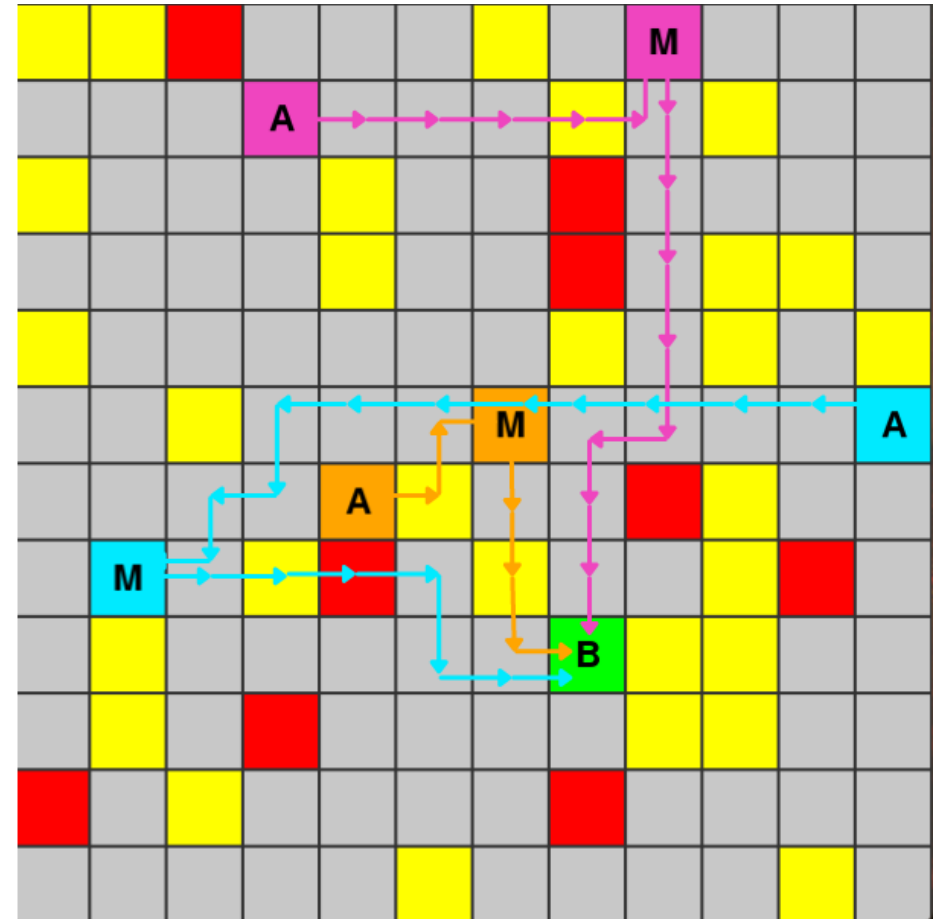
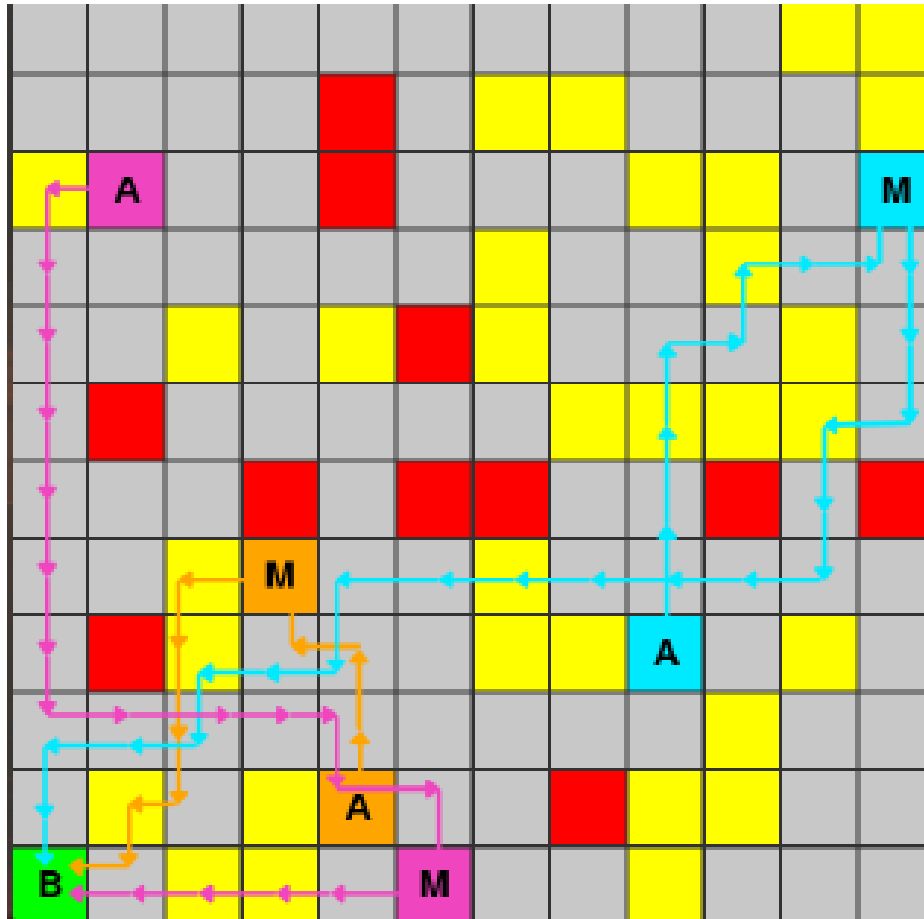
Experimentación y Pruebas Evolución Entrenamiento MAPPO



Sin Información Previa



Con Información Previa



Conclusiones y Trabajo Futuro

Conclusiones:

- MAPPO
- DDDQL
- Evitador de colisiones

Trabajo Futuro:

- Seguir entrenando
- Guiar exploración

¡Gracias!



Universidad
Internacional
de Valencia