

# **UNIVERSIDAD DE CONCEPCIÓN**

## **FACULTAD DE INGENIERÍA**



### **Departamento Ingeniería Informática y Ciencias de la Computación**

#### **Propuesta de proyecto**

**Desarrollado por:**

Francisco Cifuentes Ramírez

Marcelo Vergara Fierro

Joseph Matamala Sepúlveda

Tomás Gutiérrez Bizama



## Descripción del problema

Como plantean los autores del artículo “Optimization of On-Demand Shared Autonomous Vehicle Deployments Utilizing Reinforcement Learning” en sus conclusiones, “However, overall an RL-aided taxi dispatcher algorithm can greatly improve the performance of a deployment of SAVs by increasing the overall number of trips completed and passengers served while decreasing the wait time for passengers.”

*(Sin embargo, en general, un algoritmo de traslado en taxi potenciado por Aprendizaje por Refuerzo puede mejorar en gran manera el desempeño de la implementación de Vehículos Autónomos Compartidos, incrementando la cantidad general de viajes completados y de los pasajeros servidos, al mismo tiempo que se disminuye el tiempo de espera para los pasajeros.)*

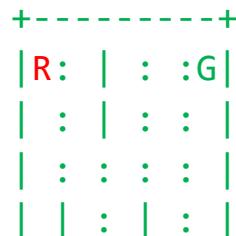
Se puede evidenciar como, en el contexto de los vehículos autónomos de taxi, existen problemas de gestión relacionados al despacho, asignación, rutas y esperas de pasajeros. Los cuales pueden solucionarse mediante un algoritmo de Aprendizaje por Refuerzo.

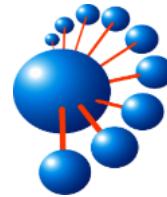
El *Taxi Problem* supone navegar pasajeros en un grid, recogerlos y dejarlos en una ubicación determinada. Existen cuatro ubicaciones designadas de recogida y de dejada. (Rojo, Verde, Amarillo y Azul) en un grid de 5x5. El taxi empieza en una casilla aleatoria y el pasajero en una de las ubicaciones descritas previamente.

La meta es mover el taxi a la posición del pasajero, recogerlo, moverlo a su destino deseado y dejarlo. Una vez el pasajero se encuentre en el destino termina el episodio.

El jugador recibe recompensas positivas por dejar al pasajero en la ubicación correcta y recompensas negativas por intentos fallidos de recogida/dejada, también por cada paso donde no se reciba ninguna recompensa.

A continuación, se puede visualizar el grid del mapa:





|Y| : |B: |  
+-----+

## Entorno del agente

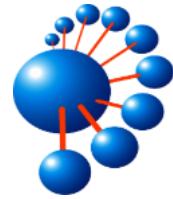
El entorno por usar es el que provee la biblioteca “*Gymnasium*” especializada en aprendizaje por refuerzo, más específicamente se usará el entorno del taxi adjunto en las referencias, las acciones son simples, todas se representan por una tupla (de formato (1,) el segundo número decide que hace) y los movimientos permitidos van desde el 0 al 3, mientras que recoger al pasajero es con el 4, y dejarlo es el 5, en cuanto a los movimientos permitidos, el 0 representa movimiento hacia abajo, el 1 movimiento hacia arriba, el 2 representa movimiento hacia la derecha y el 3 es movimiento hacia la izquierda. Este set de acciones es muy práctico ya que se trata de un espacio en 2 dimensiones, los estados iniciales están marcados por los cuadrados de colores (4) que aparecen en la simulación, estos representan donde puede aparecer tanto el pasajero como el destino de este representado por un edificio, el taxi puede aparecer en cualquiera de todas las demás casillas vacías, por último, también hay unas paredes predeterminadas que sirven de obstáculos que debe superar el agente. En total existen 500 estados posibles, pero solo son alcanzables y relevantes 400 de ellos ya que los otros 100 representan cuando el pasajero llega a su destino, y cuando el taxi esta junto al pasajero en el destino.

Las recompensas se asignan por un sistema, en el que por cada paso dado por el agente se obtiene -1 punto, a menos que ocurra otra acción con mayor recompensa, si el taxi deja correctamente al pasajero recibe +20 puntos, pero si intenta recoger o dejar al pasajero en un lugar incorrecto se le asignan -10 puntos. Además, si el agente realiza una acción inválida, como intentar moverse contra una pared, también recibe -1 punto correspondiente al paso.

## Referencias

Farama Foundation. (2025). *Taxi - Gymnasium Documentation*.

[https://gymnasium.farama.org/environments/toy\\_text/taxi/](https://gymnasium.farama.org/environments/toy_text/taxi/)



Meneses-Cime, K., Aksun Guvenc, B., & Guvenc, L. (2022). Optimization of On-Demand Shared Autonomous Vehicle Deployments Utilizing Reinforcement Learning. *Sensors*, 22(21), 8317. <https://doi.org/10.3390/s22218317>