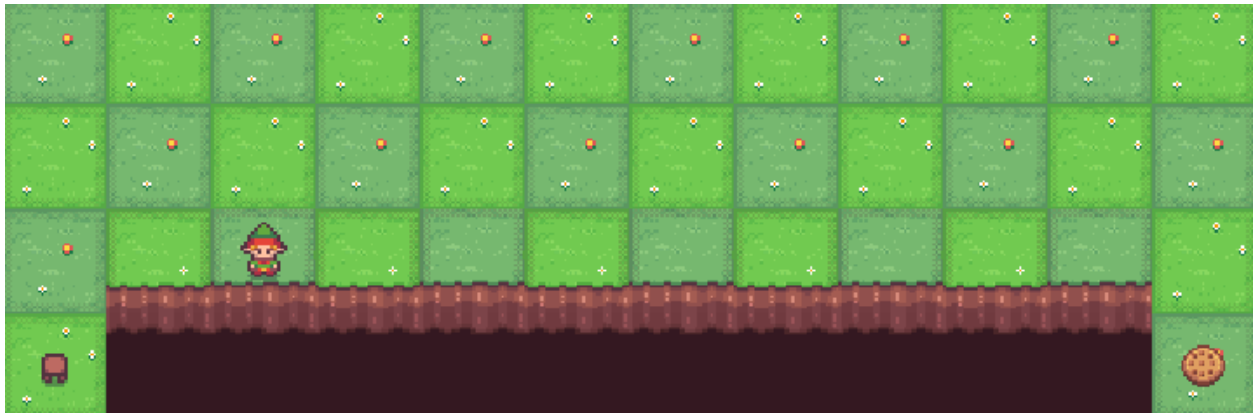


Práctico 6 - Q-Learning

El objetivo de este práctico es implementar un agente que, mediante el algoritmo de **Q-Learning**, sea capaz de navegar en un **gridworld** (mundo en forma de rejilla) hasta alcanzar un estado objetivo. El ambiente utilizado será [Cliff Walking](#) de Gymnasium.

El agente debe ser capaz de partir de un estado inicial y llegar a un estado objetivo evitando caer por el acantilado, que le penaliza severamente.



Ejemplo de estado.

Requerimientos

- 1) Implementar la clase *q_learning_agent.py* que contenga el algoritmo de Q-Learning para entrenar al agente.
- 2) Realizar al menos un entrenamiento mostrando un **plot** de la recompensa obtenida en cada episodio.
- 3) Después de entrenar al agente realizar un test en el cual, en cada instante de tiempo, el agente elija la acción óptima basada en la tabla **Q** entrenada.
- 4) Implemente un agente análogo a QLearningAgent que use el algoritmo de Double Q-Learning¹.

¹. van Hasselt, H. (2010). Double Q-learning. In Advances in Neural Information Processing Systems (NeurIPS)