**Avance Tarea 2:** Q-Learning

**Código:** EL7021-1

**Nombre:** José Luis Cádiz Sejas

**Pregunta 1:** Descripción del MDP.

* **Espacio de estados:** Dado el espacio , definimos el espacio de estados:

Donde es la función de recompensa que puede generar valores -1, 0 o NULL según si el estado es de transición, terminal o no factible respectivamente.

En particular si , estamos hablando del estado terminal:

* **Espacio de acciones**: donde {"0":"up","1":"down","2":"right","3":"left"}
* **Función de recompensa**: Función independiente de las acciones.
* **Función de transición de estados**: Solo pasa al siguiente estado si esta dentro del dominio, si no, se mantiene en el estado inicial.

**Pregunta 2:** Discretización de los estados del ambiente.

**Pregunta 3:** Código adjunto.

**Pregunta 4**: Resultados obtenidos.

**Pregunta 5**: Implementación de decaimiento lineal de épsilon. Resultados y comentarios.