

## Tarea 2 - Inteligencia Artificial 2023

### 1) (15 puntos) K-means

Utilice el algoritmo K-means, con distancia Euclidiana, para agrupar los siguientes 8 datos en 3 clusters:

A=(2,10); B=(2, 5); C=(8,4); D=(5,8); E=(7,5); F=(6,4); G=(1,2); H=(4,9);

La matriz de distancias Euclidianas es:

	A	B	C	D	E	F	G	H
A	0	$\sqrt{25}$	$\sqrt{72}$	$\sqrt{13}$	$\sqrt{50}$	$\sqrt{52}$	$\sqrt{65}$	$\sqrt{5}$
B		0	$\sqrt{37}$	$\sqrt{18}$	$\sqrt{25}$	$\sqrt{17}$	$\sqrt{10}$	$\sqrt{20}$
C			0	$\sqrt{25}$	$\sqrt{2}$	$\sqrt{4}$	$\sqrt{53}$	$\sqrt{41}$
D				0	$\sqrt{13}$	$\sqrt{17}$	$\sqrt{52}$	$\sqrt{2}$
E					0	$\sqrt{2}$	$\sqrt{45}$	$\sqrt{25}$
F						0	$\sqrt{29}$	$\sqrt{29}$
G							0	$\sqrt{58}$
H								0

Suponga que los centroides iniciales son: A, D y G. Ejecute el algoritmo k-means por 3 iteraciones. Al final de cada iteración muestre:

- Los nuevos clusters (es decir, los datos que pertenecen a cada cluster)
- Los nuevos centroides
- Dibuje en una grilla de 10x10 los puntos con su asignación en un color diferente

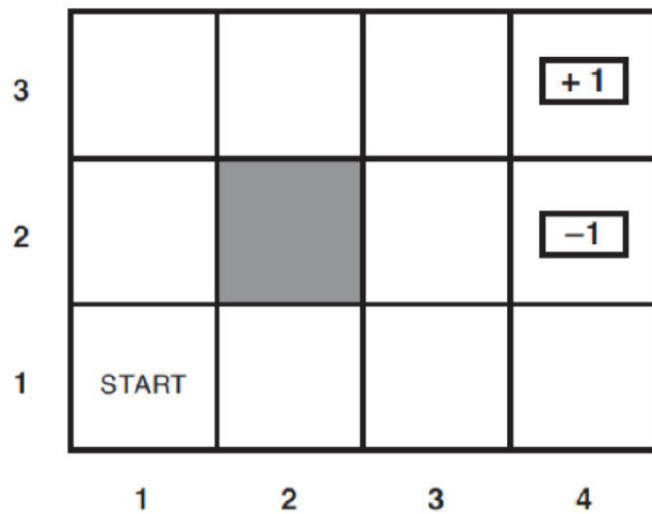
### 2) (25 puntos) Aprendizaje por refuerzo.

Adjunto a esta tarea encontrará un código en C++ llamado tutorial.cpp. Este código contiene un esqueleto de un programa que aplica técnicas de aprendizaje por refuerzo para ayudar a que un agente aprenda a moverse en diferentes ambientes. El primer ambiente se puede observar en la Figura 1, mientras que el segundo ambiente se puede apreciar en la Figura 2. Se pide implementar (o completar en) el código las instrucciones y/o funciones que permitan:

- 1.- (5 puntos) Imprimir en un archivo los pares [episodio, reward acumulado]
- 2.- (5 puntos) Implemente la estrategia (policy) epsilon-greedy para la selección de acciones (con un valor de epsilon = 0.05)

- 3.- (5 puntos) Implemente acciones estocásticas (que sólo un 80% de las veces el agente se mueva donde se le indica, el 10% de las veces se mueve a la derecha de la dirección deseada y el otro 10% de las veces se mueve a la izquierda de la dirección deseada).
- 4.- (10 puntos) Implemente y evalúe el algoritmo Q-learning en los ambientes 1 y 2. Incluya un diagrama de la curva de aprendizaje (episodio vs reward acumulado) en ambos casos.

Se debe entregar: el código completado junto con un PDF con los resultados de las 2 preguntas..



$$R(s) = -0,04$$

$$R(4,3) = +1$$

$$R(4,2) = -1$$

Figura 1

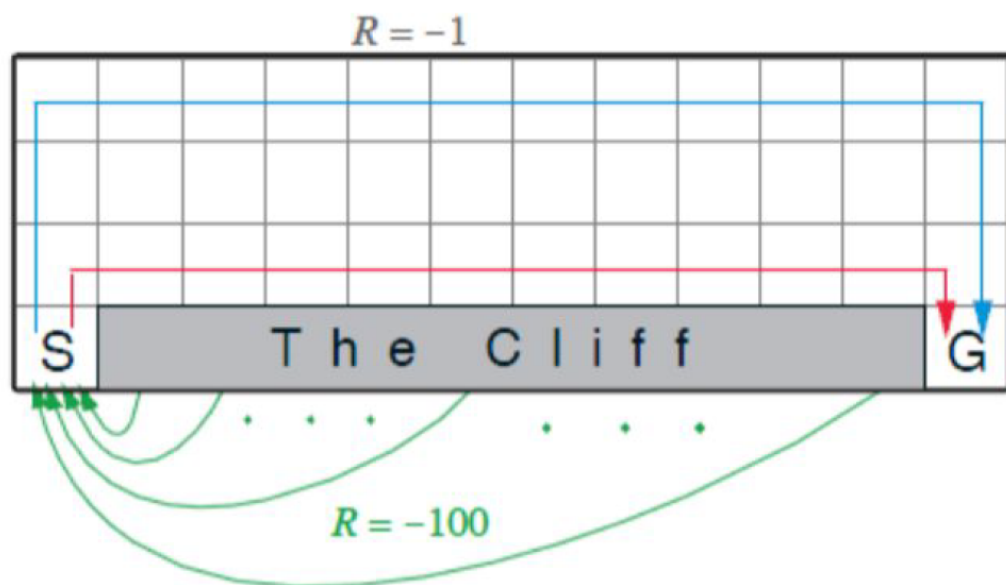


Figura 2