

# PRÁCTICA - AGENTE REACTIVO APRENDIZAJE MEDIANTE ITERACIÓN POR VALORES

## 1. PLANTEAMIENTO

Implementar el algoritmo de aprendizaje mediante iteración de valores para el entorno modelado en las siguientes matrices de transición y *reward*:

	S1	S2	S3	S4
a1	S2	S4	S1	S1
a2	S1	S3	S4	S2
a3	S2	S1	S4	S2
a4	S2	S4	S1	S3

	S1	S2	S3	S4
a1	-1	1	1	2
a2	-2	2	2	-2
a3	0	3	-2	3
a4	2	-1	4	5

El algoritmo de aprendizaje se repasa a continuación:

- Inicializar el vector  $V(s)$  arbitrariamente
- Repetir:
  - $\forall s_i \in S$  (iterar para cada  $s_i$  posible)
    - \*  $\forall a \in A$  (iterar para cada acción posible)
      - $Q(s_i, a) = R(s_i, a) + \gamma \times V\left[T(s_i, a)\right]$
    - \* Calcular  $V(s_i) = \max_a \left( Q(s_i, a) \right)$   
*Elegir el máximo valor de  $Q$  para todas las acciones posibles en el estado  $s_i$*
- hasta que los valores  $V(s)$  converjan

Deben implementar un programa que:

- (1) Lea el valor de  $\gamma$  del teclado.
- (2) Implemente el aprendizaje siguiendo el algoritmo expresado arriba.
- (3) Imprima los valores del vector  $V$  en cada estado (para ver cómo evoluciona)
- (4) Implemente la condición real de terminación, revisando que el cambio de una iteración a la siguiente sea menor a  $\epsilon = 0.001$ .

El código fuente de la inicialización de las tablas ha sido puesto en Github. **Pueden empezar de 0 si así lo prefieren.**

## 2. RESULTADO ESPERADO

### **Para el día lunes 19 de marzo, presentar el código corriendo.**

Este ejercicio de implementación es un primer paso para llegar al algoritmo de *reinforcement learning* propuesto como proyecto de primer parcial.

Sus programas deberán correr con cualquier valor de  $\gamma$  provisto, convergiendo al valor correcto (al obtenido por mi programa).

Modos de presentar:

- (1) Hacer *fork* al proyecto, e implementar las modificaciones requeridas, subiendo sus cambios a su propio proyecto (no hacer *pull requests*).
- (2) Enviando el archivo `Practica1.java` por correo. Para esta práctica no es necesario que lo envíen mediante Git, pero para el proyecto del primer parcial, será obligatorio. Vayan familiarizándose lo antes posible y para cualquier duda, estoy a su disposición.