



DEPARTAMENTO DE INFORMÁTICA  
UNIVERSIDAD CARLOS III DE MADRID

# Máster en Ciencia y Tecnología Informática

## Programación Automática Preguntas de Auto-evaluación

### Normas generales del examen

- El tiempo para realizar el examen es de **1 hora**
- No se responderá a ninguna pregunta sobre el examen
- Si se sale del aula, no se podrá volver a entrar durante el examen
- No se puede presentar el examen escrito a lápiz

1. ¿En qué consiste la propiedad de Markov?
2. En aprendizaje por refuerzo, define cómo se calcula la función de valor óptima,  $V^*(s)$ , a partir de la función de valor-acción óptima  $Q^*(s, a)$
3. Describe cómo calcular la política de comportamiento  $\pi(s, a)$  a partir de la función de valor-acción  $Q(s, a)$  del algoritmo Q-Learning
4. ¿Qué es una estrategia de exploración  $\epsilon$ -greedy?
5. ¿Qué inconvenientes tiene usar una discretización uniforme para resolver el problema de la generalización en aprendizaje por refuerzo?
6. Describe qué modificaciones es necesario hacer en el algoritmo Q-Learning para incluir la reutilización de una política pasada
7. ¿Cómo se pueden usar técnicas de aprendizaje supervisado (como regresión) para resolver el problema de la generalización en aprendizaje por refuerzo?
8. ¿Qué es el aprendizaje por demostración?
9. Describe como se puede utilizar una estrategia *Softmax* para reutilizar un conjunto de políticas en un nuevo proceso de aprendizaje por refuerzo
10. ¿Se puede hacer clustering de Procesos de Decisión de Markov?