



# **Aprendizaje Reforzado**

## **Maestría en Ciencia de Datos, DC - UBA**

Julián Martínez  
Javier Kreiner

## Repaso - Ecuaciones importantes / Métodos

$$v_{\pi}(s) := E_{\pi}[R_{t+1} + \gamma G_{t+1} | S_t = s]$$

$$q_{\pi}(s, a) := E_{\pi}[R_{t+1} + \gamma G_{t+1} | S_t = s, A_t = a]$$

$$v_{*}(s) = \max_{\pi} v_{\pi}(s)$$

$$v_{*}(s) = \max_a q_{*}(s, a)$$

$$\pi_{*}(s) = \arg \max_a q_{*}(s, a)$$



# Repaso - Programación - Soluciones

- Programación dinámica:
  - Evaluación/Predicción de política
  - Iteración de política
  - Iteración de función de valor
- Monte Carlo:
  - Predicción Monte Carlo
  - Control Monte Carlo con Políticas Epsilon Greedy