

# El problema del aprendizaje con refuerzo (I)

Lorenzo Mandow

`lawrence@lcc.uma.es`

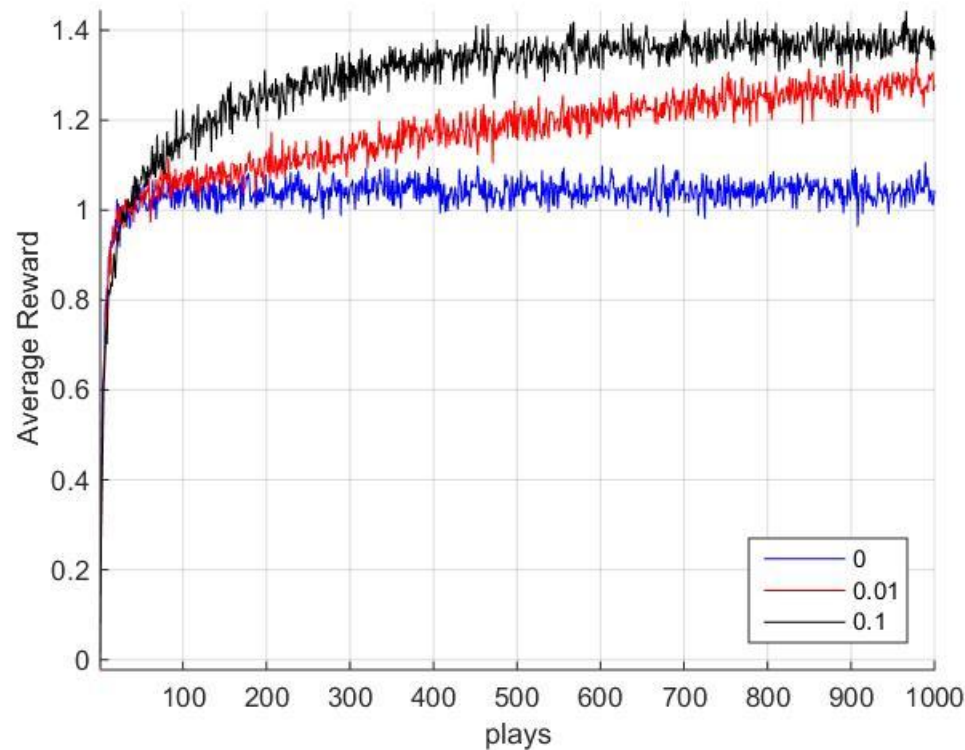
Universidad de Málaga



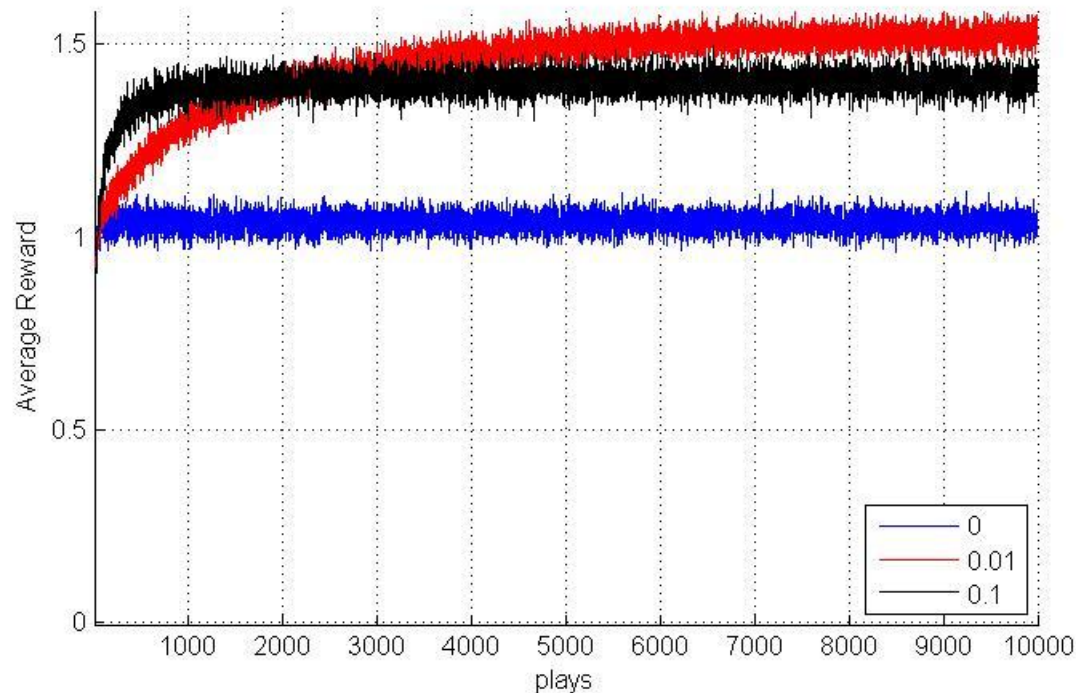
# El problema de la tragaperras

- Ejemplo: 2000 tragaperras de 10 palancas.
- Recompensas: distribución normal con varianza 1 y media obtenida de  $N(0,1)$ .
- Estimaciones: la media de las muestras
- Dos tipos de estrategias:
  - Greedy: siempre la acción con mejor estimación.
  - $\epsilon$ -greedy:
    - probabilidad  $(1 - \epsilon)$ : la acción de mejor estimación
    - Probabilidad  $\epsilon$  : elegimos aleatoriamente con distribución uniforme

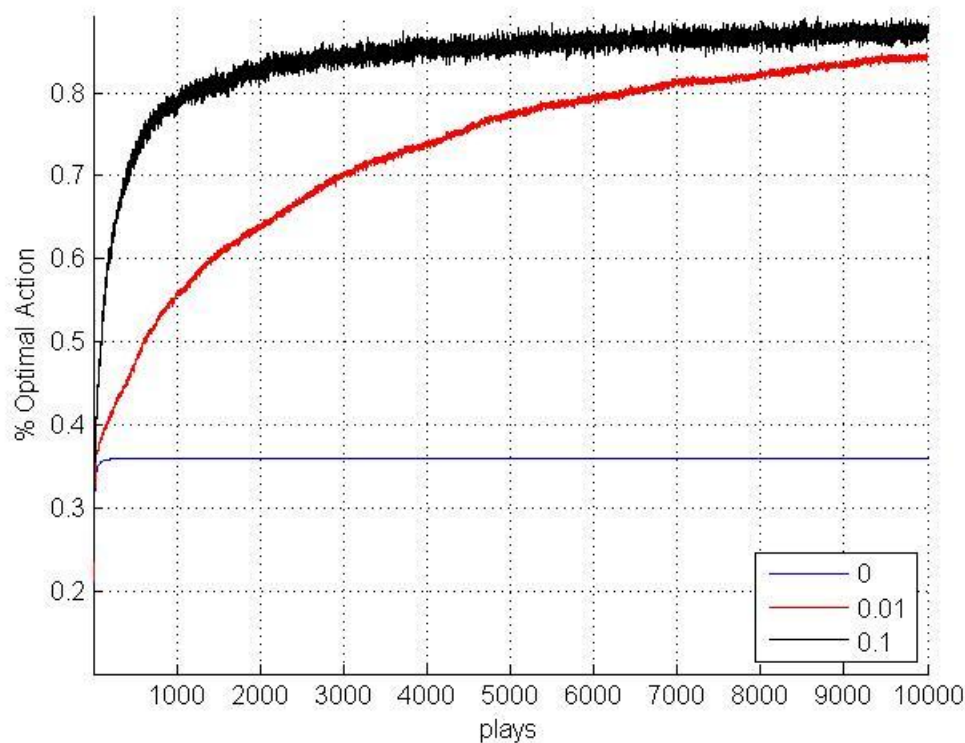
# El problema de la tragaperras



# El problema de la tragaperras



# El problema de la tragaperras



# El problema de la tragaperras

