





### Ejercicio (Valor 10 puntos)

El desafío consiste en desarrollar un script en Python que permita encontrar, mediante un algoritmo de aprendizaje por refuerzo, una solución óptima a un problema simple a elección aplicando cualquiera de las técnicas de Aprendizaje por Refuerzo del cursado, es decir, Monte Carlo ES, Monte Carlo IS (ordinario o ponderado), SARSA, Q-Learning o Deep Q-Network.

Deberá tener presente lo siguiente:

- Puede utilizar bibliotecas como Gymnasium, MiniHack, MuJoCo (u otras del interés del estudiante) para los environments, o bien crear un environment propio. No usar la biblioteca Gym dado que se encuentra obsoleta desde 2022 y debido a ello presenta diversos problemas de dependencias.
- La presentación formal debe realizarse en formato portable .pdf con un mínimo de 3 carillas que incluya lo mas relevante del código fuente. Si el documento contiene portada e índices, esas carillas no cuentan como contenido mínimo pero si suman puntos.
- Explicar en el documento en que consiste el problema y como se abordó la solución detallando inconvenientes encontrados y como se solucionó.
- Deberá contener como mínimo un gráfico de convergencia (Reward en función de los episodios). Puede usar una media móvil para suavizar la curva de convergencia de ser necesario.
- El script debe funcionar correctamente. En caso de que tenga alguna falla se le pedirá que realice la corrección. En lo posible el script debería estar diseñado para que se ejecute en Google Colab. No es obligatorio.

-  Subir este Desafío resuelto al campus de la FIUBA en formato .pdf.
-  Los grupos pueden ser de 1 integrante o más (máximo de 4 integrantes).
-  Indicar en el .pdf la URL del repositorio en donde se encuentran el código fuente.
-  Fecha límite de entrega: 13/08/2025.