

Tarea 2 - Inteligencia Artificial

Joaquin Mendez (218188)

Federico Caimi (176902)

3.b) Al visualizar gráficamente el MDP oculto del ambiente y el MDP estimado por Monte Carlo y realizando comparaciones tomando en cuenta 4 cantidades de episodios distintas, observamos:

- **Con 10 episodios:**
 - Es una cantidad insuficiente para modelar correctamente el MDP oculto por el ambiente debido a que no se toman en cuenta episodios existentes
- **Con 100 episodios:**
 - Se modela mejor, aunque siguen sin aparecer episodios o su probabilidad o recompensa no llega a ser lo suficientemente aproximada
- **Con 500 episodios:**
 - Si bien se visualizan todos los estados y sus probabilidades son bastante aproximadas en la gran mayoría de los casos no se modela la acción de **jugar** en el estado **rutina** que lleva a **fin_robo**.
- **Con 1000 episodios:**
 - Se visualizan todos los estados y sus probabilidades son bastante aproximadas en la gran mayoría de los casos se modela la acción de **jugar** en el estado **rutina** que lleva a **fin_robo**, aunque por un tema de redondeo puede ocurrir que la probabilidad de la acción jugar en el estado rutina de uno.

5.b) Al ejecutar cuatro veces el procedimiento que obtiene las recompensas medias de cada una de las políticas aplicadas al agente podemos observar que la recompensa media de del policy iteration es más estable que la de value iteration, aunque sus valores son muy similares

```
-----  
Recompensa media random policy: -21.586  
Recompensa media policy iteration: 7.873  
Recompensa media value iteration: 6.945  
  
(gym) C:\Users\feder\Desktop\IA\Tarea2>
```

```
-----  
Recompensa media random policy: -21.24  
Recompensa media policy iteration: 7.569  
Recompensa media value iteration: 8.159  
  
(gym) C:\Users\feder\Desktop\IA\Tarea2>
```

```
-----  
Recompensa media random policy: -18.365  
Recompensa media policy iteration: 7.497  
Recompensa media value iteration: 7.71  
  
(gym) C:\Users\feder\Desktop\IA\Tarea2>
```

```
-----  
Recompensa media random policy: -16.06  
Recompensa media policy iteration: 7.849  
Recompensa media value iteration: 7.508  
  
(gym) C:\Users\feder\Desktop\IA\Tarea2>
```