

IIC3675: Tarea 2

Bruno Cerda Mardini

a)

b)

c)

d)

Table 1: Resultados para GridProblem ($\gamma = 1.0$)

Tamaño Grilla	Valor Estado Inicial	Tiempo (s)
3	-8.000	0.005
4	-18.000	0.031
5	-37.333	0.175
6	-60.231	0.228
7	-93.496	0.332
8	-131.193	0.617
9	-180.001	1.005
10	-233.879	1.721

Table 2: Resultados para CookieProblem ($\gamma = 0.99$)

Tamaño Grilla	Valor Estado Inicial	Tiempo (s)
3	0.787	0.101
4	0.612	0.835
5	0.452	2.557
6	0.325	4.825
7	0.231	9.851
8	0.164	17.858
9	0.117	34.129
10	0.083	49.686

Table 3: Resultados para GamblerProblem ($\gamma = 1.0$)

Probabilidad Cara	Valor Estado Inicial	Tiempo (s)
0.25	0.067	0.218
0.40	0.284	0.309
0.55	0.612	0.435

e)

Existen problemas que toman más tiempo que otros debido al distinto tamaño del espacio de estados que tiene cada problema. El problema más lento en solucionar es el CookieProblem, en donde un estado determinado es de la forma (posicion-agente, posicion-galleta). Notar que el agente puede estar en $(\text{TamañoGrilla} * \text{TamañoGrilla})$ distintas posiciones, y la galleta también, por lo que habrá un total de $(\text{TamañoGrilla})^4$ distintos estados, lo cual escala de muy mala manera.

Luego, el GridProblem es similar al CookieProblem, con la diferencia de que el estado solo es la posición del agente, por lo que la cantidad de estados distintos es $(\text{TamañoGrilla})^2$, lo cual tampoco escala bien, pero es mejor que lo que ocurre con el CookieProblem.

Finalmente, el GamblerProblem es el más rápido de todos, principalmente debido a que solo puede tener un número de 101 estados distintos. También vale la pena mencionar que en GridProblem y CookieProblem, mientras más grande es el tamaño de la grilla, más lento es el cálculo de valores, lo cual es esperable, ya que existirán más estados y, por lo tanto, demorará más en converger. Esto se puede visualizar en los gráficos de convergencia (Figuras 1, 2 y 3).

En el GamblerProblem, mientras más uniforme sea la probabilidad (cercana a 0.5), más se demorará. Esto se debe a que el agente se quedará apostando por más tiempo, ya que perderá y ganará de manera más balanceada. Pero si la probabilidad está más cargada para cualquiera de los dos lados, entonces el tiempo es menor, ya que perderá o ganará de manera más determinista.

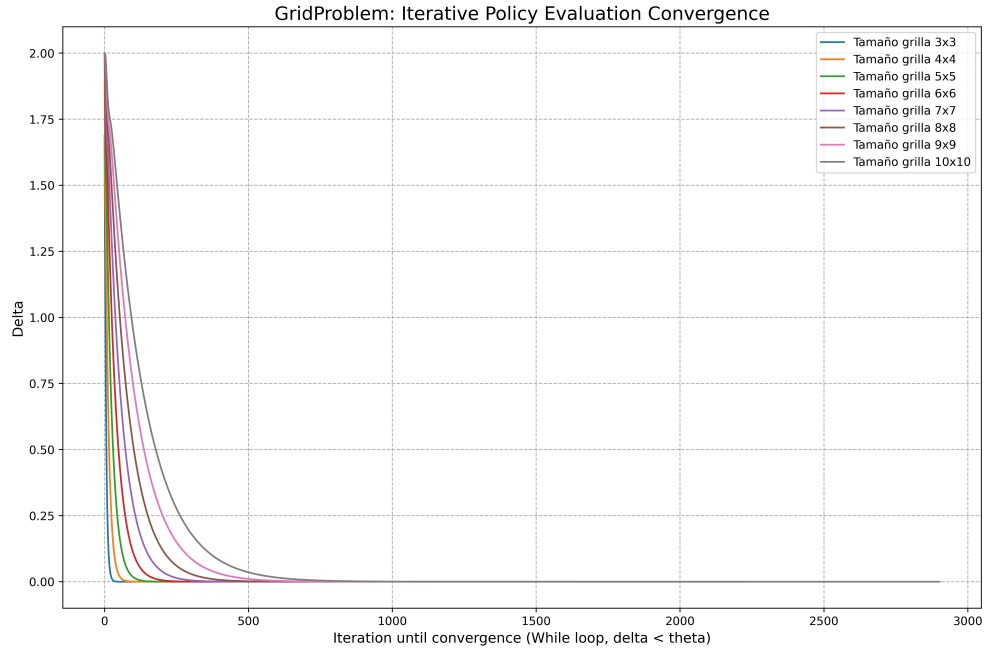


Figure 1: Convergencia para GridProblem.

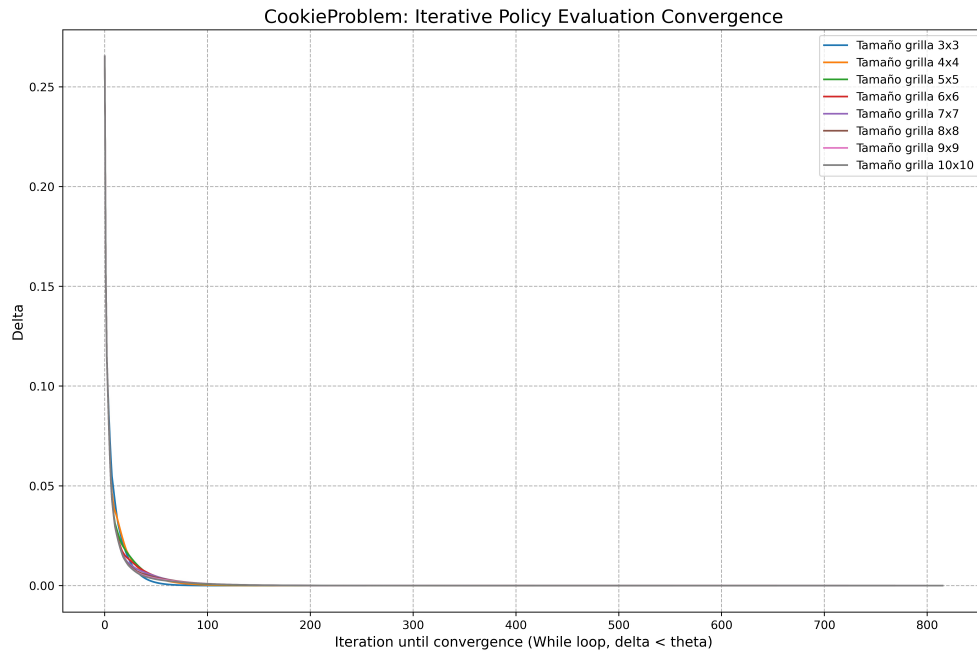


Figure 2: Convergencia para CookieProblem.

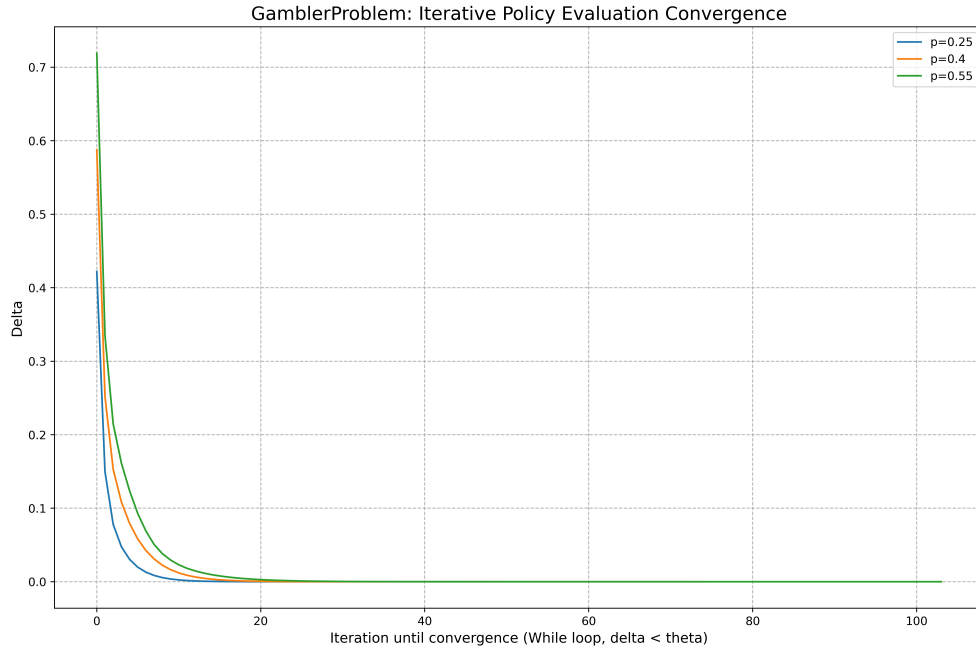


Figure 3: Convergencia para GamblerProblem.

f)

Si bajamos el valor del factor de descuento, deberíamos esperar que el tiempo de cómputo sea menor, ya que los valores deberían converger de manera más rápida. Probé todos los problemas con un $\gamma = 0.5$ y las soluciones se encuentran mucho más rápido.

Esto ocurre ya que con un gamma pequeño, las recompensas serán más relevantes en la fórmula de Bellman, por lo que los valores de los estados estarán más ligados a las recompensas inmediatas que a las recompensas futuras. De esta manera, el vínculo entre los valores de los estados se reduce, por lo que la convergencia ocurre más rápido.

g)

h)

i)

j)

k)

l)

m)

n)