

Problema del Robot

Jazmin Flores e Irasema Pedroza

2025-09-10

Table of contents

Contexto

3

Contexto

Tenemos un robot en cual tiene dos estados: bateria alta y bateria baja, estos estados sera denotados como high y low respectivamente. Las acciones posibles para el estado low son: buscar, esperar y recargar. En cambio los del estado high son buscar y esperar. El resumen de las probabilidades del robot estan dadas en la siguiente tabla:

Table 1: Demonstration of pipe table syntax

s	a	s'	p(s' s, a)	r(s, a, s')
high	search	high	$\alpha = 0.5$	r_search=3
high	search	low	$1 - \alpha = 0.5$	r_search= -3
low	search	high	$1 - \beta = 0.5$	r_search =3
low	search	low	$\beta = 0.5$	r_search=3
high	wait	high	1	r_wait =2
high	wait	low	0	—
low	wait	high	0	—
low	wait	low	1	r_wait =2
low	recharge	high	1	r_wait=2
low	recharge	low	0	—

$$\begin{aligned}
v_*(h) &= \max \left\{ \begin{aligned} &p(h | h, s) [r(h, s, h) + \gamma v_*(h)] + p(l | h, s) [r(h, s, l) + \gamma v_*(l)], \\ &p(h | h, w) [r(h, w, h) + \gamma v_*(h)] + p(l | h, w) [r(h, w, l) + \gamma v_*(l)] \end{aligned} \right\} \\
&= \max \left\{ \begin{aligned} &\alpha [r_s + \gamma v_*(h)] + (1 - \alpha) [r_s + \gamma v_*(l)], \\ &1 [r_w + \gamma v_*(h)] + 0 [r_w + \gamma v_*(l)] \end{aligned} \right\} \\
&= \max \left\{ \begin{aligned} &r_s + \gamma [\alpha v_*(h) + (1 - \alpha) v_*(l)], \\ &r_w + \gamma v_*(h) \end{aligned} \right\} \\
&= \max \left\{ \begin{aligned} &3 + 0.5 [0.5 v_*(h) + (0.5) v_*(l)], \\ &2 + 0.5 v_*(h) \end{aligned} \right\}
\end{aligned}$$

$$= \max \left\{ \begin{array}{l} 3 + 0.5 [0.5v_*(h) + (0.5)v_*(l)], \\ 4 \end{array} \right\}$$

de igual forma tenemos $v_*(l)$ esta dado por:

$$v_*(l) = \max\{0.5[0.5v_*(h) + 0.5v_*(l)], 2 + 0.5v_*(l), 0.5v_*(h)\}$$

supongamos que el maximo de $v_*(h)$ es $2 + 0.5v_*(h)$ despejamos y el resultado es $v_*(h) = 4$,
apartir de este valor podemos obtener $v_*(l)$ el cual es:

$$v_*(l) = \max\{1 + 0.5v_*(l), \\ 2 + 0.5v_*(l), \\ 2\}$$

Observemos que los unicos candidados son: $2 + 0.5v_*(l)$ y 2 pero en el caso que si $v_*(l) < 0$
tendriamos que $v_*(l) = 2$ pero como 2 no es un numero negativo tenemos que si $v_*(h) = 4$
entonces $v_*(l) = 2 + 0.5v_*(l)$

Ahora supondremos que $v_*(h) = 3 + 0.5 [0.5v_*(h) + (0.5)v_*(l)]$

Si $v_*(h) = 3 + 0.25v_*(h) + 0.25v_*(l)$, entonces $v_*(h) = 3 + 0.25v_*(l)$

por lo tanto para el estado low tenemos que

$$v_*(l) = \max\{0.25(3 + 0.25v_*(l)) + 0.25v_*(l), \\ 2 + 0.5v_*(l), \\ 0.5(3 + 0.25v_*(l))\}$$