Q4 - Micro Black Jack

Value Iteration

⇒ set $V(s) = 0$; $\forall s$

⇒ use Bellmans equation until values converge

$$V(s) = \max\left(R(s, \text{stop}, \text{done}) \sum_{s'} P(s' | s, \text{draw})\, V(s')\right)$$

Iteration 1; Initial guess $[V(s) = 0]$

State (s)	Stop Value R(s, stop)	Draw Value $[\frac{1}{3}V(s+2) + \frac{1}{3}V(s+3) + \frac{1}{3}V(s+4)$	V(s)*
0	0	$\frac{1}{3}(2+3+4) = 3$	3
2	2	$\frac{1}{3}(4+5+6) = 5$	5
3	3	$\frac{1}{3}(5+6+0) = 3.66\overline{6} \approx 3.67$	3.67
4	4	$\frac{1}{3}(6+0+0) = 2$	4
5	5	0	5
6	6	0	6

Update Policy

V(s)

↳ if $V(s, \text{draw}) > V(s, \text{stop}) =$ Draw

else stop

Old policy

(state, policy(d/s), V(s, draw), V(s, stop))

{(0, d, 3, 0),
(2, d, 5, 2),
(3, d, 3.67, 3),
(4, d, 2, 4),
(5, s, 0, 5),
(6, s, 0, 6)}

New policy

(state, policy(d/s), V(s, draw), V(s, stop))

{(0, d, 3, 0),
(2, d, 5, 2),
(3, d, 3.67, 3),
(4, s, 2, 4),
(5, s, 0, 5),
(6, s, 0, 6)}

Iterate until convergence

State (s)	Stop Value R(s, stop)	Draw Value $[\frac{1}{3}V(s+2) + \frac{1}{3}V(s+3) + \frac{1}{3}V(s+4)$	V(s)*
0	0	$\frac{1}{3}(5+3.67+4) = 4.22$	4.22
2	2	$\frac{1}{3}(4+5+6) = 5$	5
3	3	$\frac{1}{3}(5+6+0) = \frac{11}{3} = 3.67$	3.67
4	4	2	4
5	5	0	5
6	6	0	6