Q3 ~ Micro Black Jack

States: $S = \{0, 2, 3, 4, 5, \text{Done}\}$

Actions: $A = \{\text{Draw}, \text{Stop}\}$

Transition P: $P(s' | s, a)$ [MDP]

⇒ If "draw" is chosen
- ~ $P(s+2 | s, \text{draw}) = 1/3$
- ~ $P(s+3 | s, \text{draw}) = 1/3$
- ~ $P(s+4 | s, \text{draw}) = 1/3$

⇒ if "stop" is chosen
- ~ game moves to DONE (terminal)

Reward Function: $R(s, a, s')$

⇒ if "stop" is chosen
- ~ $R(s, \text{stop}, \text{done}) = s$; where s is the score @ stopping point

⇒ if "draw" is chosen
- ~ $R(s, \text{Draw}, s') = 0$; no immediate reward

Discount Factor: $\gamma = 1$ (game ends quickly, no γ needed)

Value Iteration

⇒ set $V(s) = 0$; $\forall s$

⇒ use Bellmans equation until values converge

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

Tables + Iterations

Computing $V_1(s)$

State s	Action	Value Calculation	$V_1(s)$
0	Stop Draw	0 $\frac{1}{3}(0 + 1V_0(2)) + \frac{1}{3}(0 + 1V_0(3)) + \frac{1}{3}(0 + 1V_0(4)) = 0$	0 0
2	Stop Draw	2 $\frac{1}{3}(0 + 1V_0(4)) + \frac{1}{3}(0 + 1V_0(5)) + \frac{1}{3}(0 + 1V_0(0)) = 0$	2 2
3	Stop Draw	3 $\frac{1}{3}(0 + 1V_0(5)) + \frac{2}{3}(0 + 1V_0(0)) = 0$	3 3
4	Stop Draw	4 0 (because score ≥ 6 is a loss)	4 4
5	Stop Draw	5 0 (loss)	5 5

$\Rightarrow$ Observations: s_4 & s_5 have max values when stopping $\Rightarrow$ Drawing = loss

Computing $V_2(s)$

State s	Action	Value Calculation	$V_1(s)$
0	Stop Draw	0 $\frac{1}{3}(2) + \frac{1}{3}(3) + \frac{1}{3}(4) = 3$	0 3
2	Stop Draw	2 $\frac{1}{3}(4) + \frac{1}{3}(5) + \frac{1}{3}(0) = 3$	2 3
3	Stop Draw	3 $\frac{1}{3}(5) + \frac{2}{3}(0) = \frac{5}{3} + 0 = \frac{5}{3}$	3 3

$\Rightarrow$ Observations: s_0 has an expected value of 3 if drawing is chosen

Computing $V_3(s)$

State s	Action	Value Calculation	$V_1(s)$
0	Stop Draw	0 $\frac{1}{3}(3) + \frac{1}{3}(3) + \frac{1}{3}(4) = \frac{10}{3}$	0 $\frac{10}{3}$

$\Rightarrow$ Observations: s_0 reaches a stable expected value when drawing

Final Values + Policy Extraction

State, Action

$$T = \{ (0, \text{Draw}), (2, \text{Draw}), (3, \text{Stop}), (4, \text{Stop}), (5, \text{Stop}); (S_n, A) \}$$