PAGULAYAN, KAMIRA ALLISON F. COM221

EPISODE	STEP	STATE	ACTION	REWARD	NEXT STATE
EPISODE 1	1	13, 11, T	HIT	0	19, 11, T
	2	19, 11, T	STAND	0	19, 21, T
	3	19, 21, T	STAND	-1	BUST
EPISODE 2	1	18, 2, F	HIT	0	22, 2, F
	2	22, 2, F	STAND	0	22, 2, F
	3	22, 2, F	STAND	0	22, 27, F
	4	22, 27, F	STAND	-1	BUST
EPISODE 3	1	19, 11, T	HIT	0	23, 11, T
	2	23, 11, T	STAND	0	23, 21, T
	3	23, 21, T	STAND	-1	BUST
EPISODE 4	1	15, 4, F	HIT	0	17, 4, F
	2	17, 4, F	HIT	0	27, 4, F
	3	27, 4, F	STAND	0	27, 19, F
	4	27, 19, F	STAND	-1	BUST
EPISODE 5	1	19, 6, F	HIT	0	25, 6, F
	2	25, 6, F	STAND	0	25, 6, F
	3	25, 16, F	STAND	-1	BUST
EPISODE 6	1	12, 3, F	HIT	0	19, 3, F
	2	19, 3, F	STAND	0	19, 20, F
	3	19, 20, F	STAND	-1	BUST
EPISODE 7	1	14, 6, F	HIT	0	19, 6, F
	2	19, 6, F	STAND	0	19, 26, F
	3	19, 26, F	STAN	1	WIN
EPISODE 8	1	6, 10, F	HIT	0	16, 10, F
	2	16, 10, F	STAND	0	16, 22, F
	3	16, 22, F	STAND	1	WIN
EPISODE 9	1	18, 2, F	STAND	0	18, 2, F
	2	18, 2, F	STAND	0	18, 22, F
	3	18, 22, F	STAND	1	WIN
EPISODE 10	1	10, 5, F	HIT	0	15, 5, F
	2	15, 5, F	STAND	0	15, 20, F
	3	15, 20, F	STAND	-1	BUST

MONTE CARLO

STATE	RETURN G	N(S)	Old V(s)	New V(s)
14, 6, F	+1	1	0	1
19, 6, F	+1	1	0	1
19, 26, F	+1	1	0	1

TEMPORAL DIFFERENCE UPDATE

STEP	STATE	REWARD	NEXT STATE S'	OLD V(S)	NEW V(S)
1	(14, 6, F)	0	(19, 6, F)	0	0
2	(19, 6, F)	0	(19, 26, F)	0	0
3	(19, 26, F)	+1	NONE / WIN	0	0.5

PREVIOUSLY VISITED STATE

19, 6, F	0	19, 26, F	0	0.25
14, 6, F	0	19, 6, F	0	0.125

SOLUTION:

$$V(s_t) \leftarrow V(s_t) + a\,[R_t + 1 + \gamma V(s_{t+1}) - V(s_t)]$$

1) $V(19, 26, F) = 0 + 0.5(1 + 0 - 0) = 0.5$

2) $V(19, 6, F) = 0 + 0.5(0 + 0.5 - 0) = 0.25$

3) $V(14, 6, F) = 0 + 0.5(0 + 0.25 - 0) = 0.125$

$$V(s_t) \leftarrow V(s_t) + a\,[G_t - V(s_t)]$$

1) $V(19, 26, F) = 0 + 0.5(1 - 0) = 0.5$

2) $V(19, 6, F) = 0 + 0.5(1 - 0) = 0.5$

3) $V(14, 6, F) = 0 + 0.5(1 - 0) = 0.5$