NO.:
DATE:

aldrin A. Balucay COM-221

Step	State	Action	Reward	Next State
1 2	17,10 F	Hit	0	27,10 F
2	27,10 F	~~Hit~~ Stand	-1	BUST ~~(27 T to F)~~
1 2 3	21,6 True	Stand	1	21,6 T W
~~1 2 3~~	21,2 True	Stand	1	21,2 T W
1 2 3	9,8 F	Hit	0	19,8 F
	19,8 F	Hit	0	29,8 F
	29,8 F	Stand	-1	BUST
1 2	17,5 F	Hit	0	22,5 F
	22,5 F	Stand	-1	BUST
1 2 3	13,3 F	Hit	0	18,3 F
	18,3 F	Hit	0	28,3 F
	28,3 F	Stand	-1	BUST
1 2 3	12,5 True	Hit		20,5 T
	20,15 True	Stand ~~Stand~~	0	20,15 T
	20,17 True	Stand	1	20,17 T W
1 2 3 4	13,4 F	Hit	0	21,6 F
	21,6 F	Stand	0	21,15 F
	~~21,17~~ 15 F	Stand	0	21,17 F
	21,17 F	Stand	1	21,17 F W
1 2	18,7 F	Hit	0	28,7 F
	28,7 F	Stand	-1	BUST
1 2 3	21,10 F	Stand	1	Win

Aldwin A. Ramirez

Monte Carlo

(17, 5, F)

(22, 5, F)

Step	State	Reward	Next State	old V	New V
1	(17, 5, F)	0	(22, 5, F)	0	-0.5
2	(22, 5, F)	-1	BUST	-1	-0.75
3	(22, 5, F)	1	BUST	-0.5	

Step: 1

a State = (17, 5, F) =

b. N (17, 5, F) = 0

c. V (17, 5, F) = 0

d. N (17, 5 F) = 1

e V (17, 5, F) = 0 + $\frac{1}{1}$ (-1 - 0) = -1

Step 2:

a. = (22, 5, F)

b. N(22, 5, F) = 0

c = V(22, 5, F) = 0

d = N(22, 5, F) = 1

e = V(22, 5, F) = 0 + $\frac{1}{1}$ (-1 - 0) = -1

State	Return	N(s)	Old(V)	New V(s)
17, 5, F	0	1	0	-1
22, 5, F	-1	1	0	-1

Temporal

a. State = (17, 5, F)

b Reward = 0

c. Next State = (22, 5, F)

d. Update: (17, 5, F) = 0 + 0.5(0 + V(22, 5, F) - 0) = 0

Step: 3

V(s) = 0.05(-1 + (-0.5) + 0) = -0.75

Step 2

a. State = (22, 5, F)

b. Reward = r = -1

. Next State = Bust

Update: V(17, 5, F) = 0 + 0.5 (0 + 0 - 1) = -0.5

d. Update: V(22, 5, F) = 0 + 0.5(-1 + 0 - 0) = -0.5

[-1 + 0] = -0.5 × 1.5 = -0.75