Christian Joshua Alberto

Ep 1

Step	State	Action	Reward	Next state
1	17,4,F	Stand	0	17,17 F
2	17,17,F	Stand	0	DRAW

Ep 2

Step	State	Action	Reward	Next State
1	18,9,F	Stand	0	18,18,F
2	18,18,F	Stand	0	~~18,18,F~~ DRAW

Ep 3

Step	State	Action	Reward	Next state
1	16,2,F	HIT	0	23,2,F
2	23,2F	Stand	-1	BUST

Ep 4

Step	State	Action	Reward	Next state
1	20,2,F	Stand	0	20,23,F
2	20,23,F	Stand	+1	WIN

Ep 5

Step	State	Action	Reward	Next State
1	14,10,F	Hit	0	19,10,F
2	19,10,F	Stand	0	19,21,F
3	19,21,F	Stand	-1	LOST

Ep 6

Step	State	Action	Reward	Next State
1	20,6,F	Stand	0	20,28,F
2	20,28,F	Stand	+1	WIN

Ep 7

Step	State	Action	Reward	Next State
1	19,3,T	Stand	0	19,23,T
2	19,23,T	Stand	+1	WIN

Ep 8

Step	State	Action	Reward	Next Episode
1	17,10,F	Stand	0	17,19,F
2	17,19,F	Stand	-1	LOST

Ep 9

Step	State	Action	Reward	Next state
1	20,9,F	Stand	0	20,19,F
2	20,19,F	Stand	+1	WIN

Ep 10

Step	State	Action	Reward	Next State
1	14,10,F	Hit	0	22,10,F
2	22,10,F	Stand	-1	BUST

Part B: Monte Carlo Update

Ep 4:

State	Return G	Visit Count	Old V(s)	New V(s)
20,2,F	+1	1	0	1
20,23,F	+1	1	0	1

State 1 : New V(s)

$$0 + \frac{1}{1}(1-0) = 1$$

State 2 New V(s)

$$0 + \frac{1}{1}(1-0) = 1$$

Part C: Temporal Difference(0) Update

Step 1:

a. State = $(20,2,F)$

b. Reward $r = 0$

c. Next State = $(20,23,F)$

d. Update ⇒ $V(20,2,F) = 0 + 0.5 (0 + 0 - 0) = 0$

Step	State	Reward	Next State	Old V(s)	New V(s)
1	(20,2,F)	0	20,23,F	0	0
2	—	—	—	—	—

Update : $V(20,23,F) = 0 + 0.5(1+0-0) = 0.5$

Step 2:

Step	State	Reward	Next State	Old	New
1	(20,2,F)	0	20,23,F	0	0
2	(20,23,F)	+1	WIN	0	0.5

Step 3:

Update : $V(20,2,F) = 0 + 0.5(1 + 0.5 - 0) = 0.25$

Step	State	Reward	Next State	Old	New
1	20,2,F	0	20,23,F	0	0.25
2	20,23,F	+1	WIN	0	+0.5