Balder, John Vernon D

NO.:
DATE:

| Step | State | Action | Reward | Next State | |
|---|---|---|---|---|---|
| 1 | 12, 5, F | Hit | 0 | 22, 5, F | > ep 1 |
| 2 | 22, 5, F | St | -1 | ~~22~~ Bust | |
| 3 | 19, 5, F | St | 0 | 19, 5, F | > ep 2 |
| 4 | 19, 25, F | St | 1 | WIN | |
| 5 | 17, 3, F | St | 0 | 17, 3, F | > ep 3 |
| 6 | 17, 22, F | St | 1 | WIN | |
| 7 | 20, 5, F | St | 0 | 20, 5, F | > ep 4 |
| 8 | 20, 23, F | St | 1 | WIN | |
| 9 | 16, 4, F | Hit | 0 | 20, 4, F | > ep 5 |
| 10 | 20, 4, F | St | | Bust | |
| 11 | 11, 8, F | Hit | 0 | 21, 8, F | > ep 6 |
| 12 | 21, 8, F | St | 1 | 21, 23, F WIN | |
| 13 | 20, 10, F | St | 0 | 20, 14, F | > ep 7 |
| 14 | 20, 14, F | St | 1 | 20, 30, F WIN | |
| 15 | 6, 10, F | Hit | 0 | 20, 10, F | > ep 8 |
| 16 | 20, 10, F | St | +1 | 20, 22, F WIN | |
| 17 | 11, 8, F | Hit | 0 | 15, 8, F | ep 9 |
| 18 | 15, 8, F | Hit | 0 | 18, 8, F | |
| 19 | 18, 8, F | Hit | 0 | 21, 8, F | |
| 20 | 21, 8, F | St | 1 | 27, 27, F WIN | |
| 21 | 17, 7, F | Hit | 0 | 23, 7, F | > ep 10 |
| 22 | 23, 7, F | St | | 23, 9, F Lose | |

NO.:
DATE:

Monte carlo (episode 6)

| State | Return | N(s) | Old V(s) | New V(s) |
|---|---|---|---|---|
| (11, 8, F) | 1 | 1 | 0 | 1 |
| (21, 8, F) | 1 | 1 | 0 | 1 |
| (21, 23, F) | 1 | 1 | 0 | 1 |

V(11, 8, F) = 0

N (11, 8, F) = 1

update : $V(11,8,F) = 0 + \frac{1}{1}(1-0) = 1$

V(21, 8, F) = 0

N(21, 8, F) = 1

up : $V(21,8,F) = 0 + \frac{1}{1}(1-0) = 1$

V(21, 23, F) = 0

N(21, 23, F) = 1

$V(21,23,F) = 0 + \frac{1}{1}(1-0) = 1$

Temporal

| State | Reward (Return) | Next State (N(s)) | Old V(s) | New V(s) |
|---|---|---|---|---|
| (11, 8, F) | 0 | (21, 8, F) | 0 | 0 |
| (21, 8, F) | 0 | (21, 23, F) | 0 | +0.5 |
| (21, 23, F) | 1 | WIN | | |

(21, 23, F) 1 WIN

V(11, 9, F)

r = 0

Next : (19, 10, F)

up: $V(11,8,F) = 0 + 0.5(0 + 0 - 0) = 0$

V(21, 8, F)

r = 1

Net: (21, 23, F)

up: $V(21,23,F) = 0 + 0.5(1 + 0 - 0) = 0.5$

$V(11,8,F)$ ; r = 1

up: $V(11,8F) = 0 + 0.5(1 + 0.5 - 0) = 0.25$

| State | R | Next | Old | New |
|---|---|---|---|---|
| 11, 8, F | 1 | 21, 8, F | 0 | 0.25 |
| 21, 8, F | 1 | 21, 23, F | 0 | 0.5 |

NO.:
DATE:

Temporal

| State | Reward | Next | Old | New |
|---|---|---|---|---|
| (11, 8, F) | 0 | (21, 8, F) | 0 | 0 / 0.125 |
| (21, 8, F) | 0 | (21, 23, F) | 0 | 0 / 0.25 |
| (21, 23, F) | 1 | (win) | 0 | 0.5 |

V(11, 8, F)
r = 0
Next : (21, 8, F)
update : V(11, 8, F) = 0 + 0.5 (0 + 0 - 0) = 0

V(21, 8, F)
r = 0
next : (21, 23, F)
up : V(21, 23, F) = 0 + 0.5 (0 + 0 - 0) = 0

V(21, 23, F)
r = 1
next : win
update : V(21, 23, F) = 0 + 0.5 (1 + 0 - 0) = 0.5

V(11, 8, F)
update : V(11, 8, F) = 0 + 0.5 (0 + 0 - 0) = 0

V(21, 8, F)
up : V(21, 8, F) = 0 + 0.5 (0 + 0.5) = 0.25

V(11, 8, F)
update : V(11, 8, F) = 0 + 0.5 (0 + 0.25 - 0) = 0.125