# **RL** Assignment-2

Ashish Sethi MT-18024

Question - 2	3
Results	3
Initial Value Function	3
Value Function after policy iteration	3
Question - 4	4
Results	4
Initial Value Function	4
Value Function after policy iteration	4
Question - 6	5
Policy Improvment	5
Intial Random Policy	5
Final Policy	5
Value Iteration	6
Intial Policy	6
Final Policy	6

## Question - 2

#### Results

Initial Value Function

0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

Value Function after policy iteration

3.3	8.8	4.4	5.3	1.5
1.5	3.0	2.3	1.9	0.5
0.1	0.7	0.7	0.7	-0.4
-1	-0.4	-0.4	-0.6	-1.2
-1.9	-1.3	-1.2	-1.4	-2.0

## Question - 4

#### Results

Initial Value Function

0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

Value Function after policy iteration

22	24.4	22.	19.4	17.5
19.8	22.	19.8	17.8	16
17.8	19.8	17.8	16	14.4
16	17.8	16	14.4	13.0
14.4	16.0	14.4	13.0	11.7

## Question - 6

# Policy Improvment

Intial Random Policy

3	2	3	3
2	2	3	2
0	1	2	3
3	0	0	0

## Final Policy

0	0	0	0
1	0	0	3
1	0	2	3
1	2	2	0

### Value Iteration

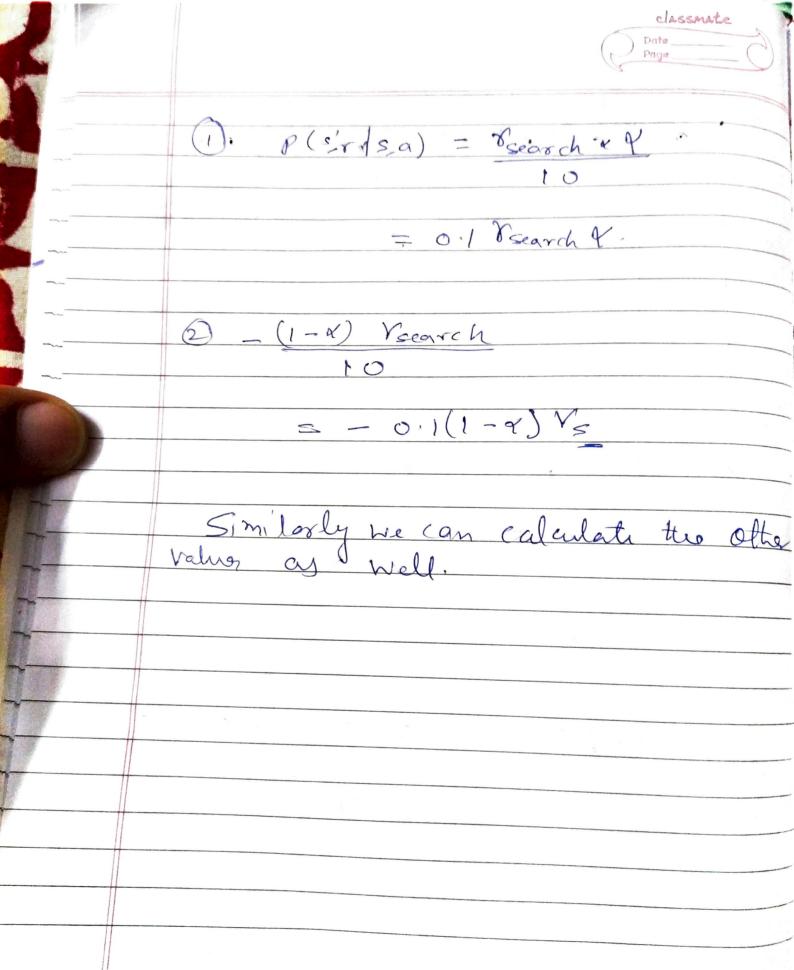
Intial Policy

0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

Final Policy

0	0	0	0
1	0	0	3
1	0	2	3
1	2	2	0

RL Assignment-2 classmate · Ashish seth? Page . MT 18024 Question-1 a p(s'|sa) | \* (sasi) 5 **s** / P(dr | 5a) hight high Sparch 0.1 of rsearch Ysearch Ligh Search low 1-4 -0.1 (1-9) Franch rseorch low Search high 1-13 -3 low search low Osearch -0.13 rs high Wart high Twain 0.1 In high wait low 0 low wait high low Wort low Thoran f 0.120 low recharge high low socher & Jow As we know that  $r(sa,s') = \sum_{i=1}^{n} r p(s',r|s,a)$ rer p(s/s\_a)  $p(s'|s|s|a) = \gamma(s|a|s') p(s|s|a)$ E Y





93 Exercise 3.15

we know that

V \* (s) = ET SR St = 3 }

- E S S 7 | S = 5 ]

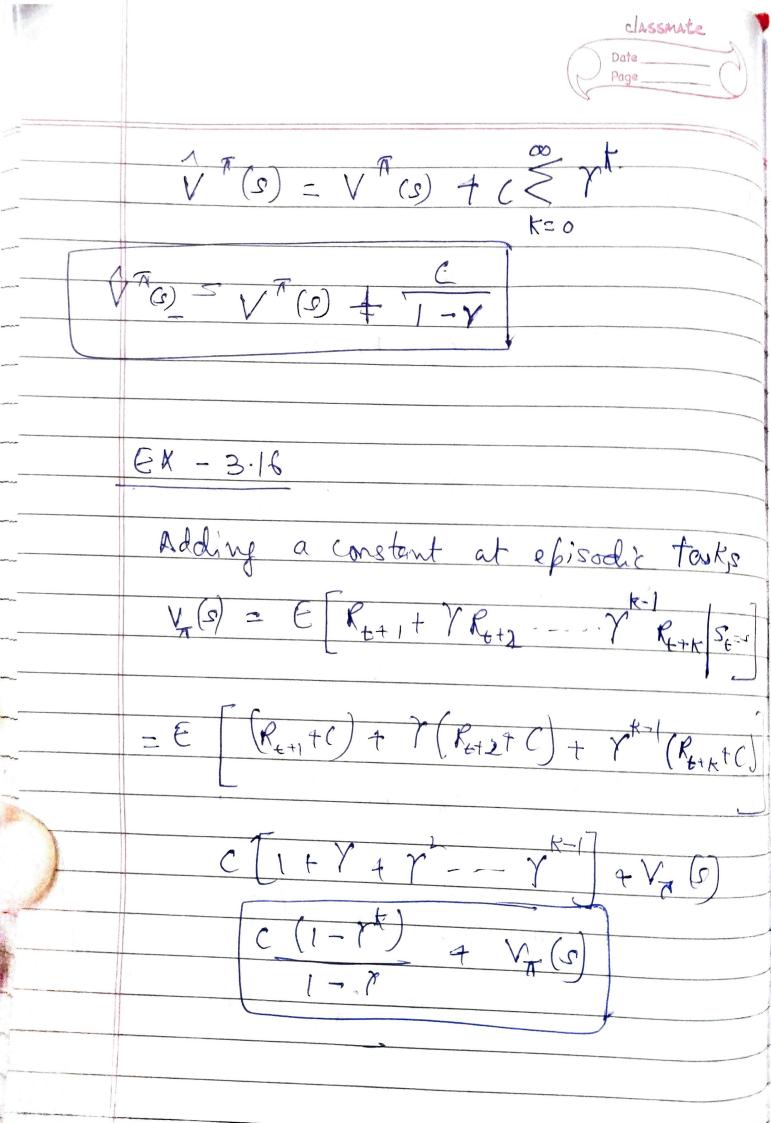
Add constant in reword.

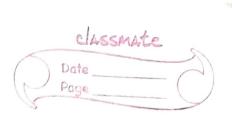
Potker = Petker + C

 $V^{r}(s) = E_{r} \left[ \sum_{k=0}^{\infty} Y^{r} Y_{t+k+1} \right] S_{0} = S_{0}^{r}$ 

= ET S & r r r | S = sy

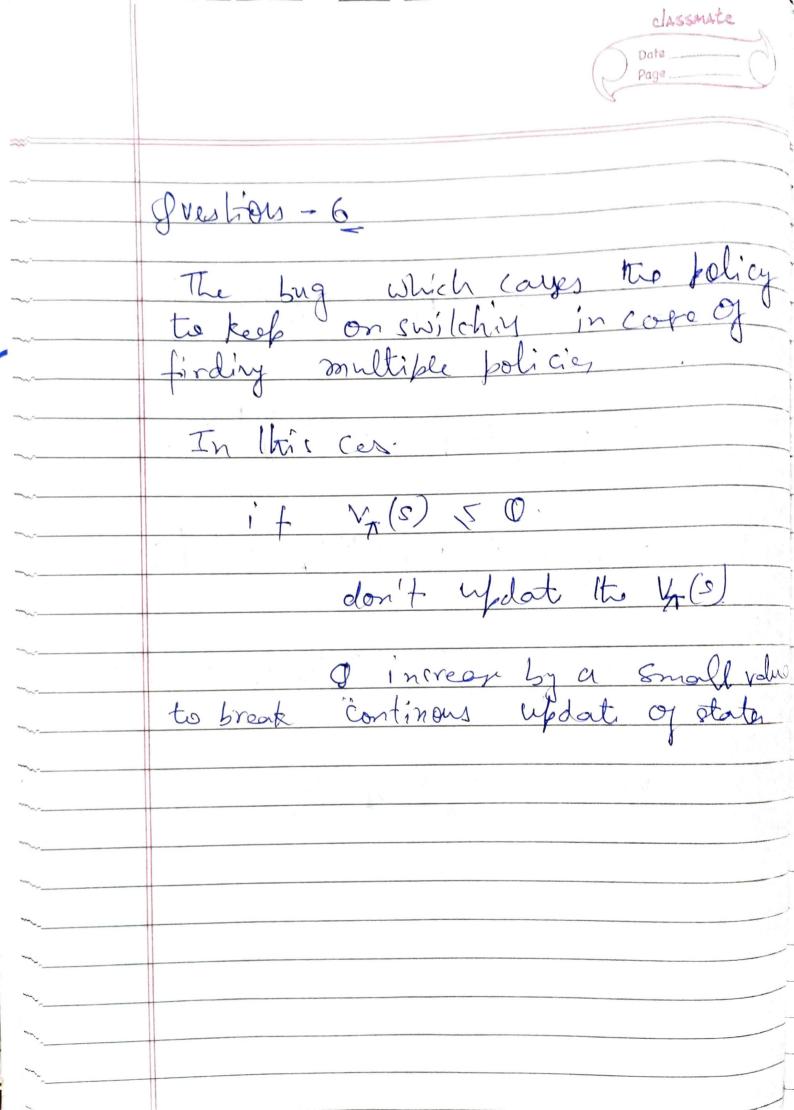
+ Fr Extc Stesy





grestion -5

$$V_{r}(s) = \mathcal{E}_{r}\left[R_{t+1} + \gamma V_{r}(s_{t+1})\right] S_{t} = s$$



classmate.

Question 4.

Non linear Solution of bellmon equation using linear programming

 $V(s) \geq R(s) + \gamma \max_{q \in A} \leq P(s, |s, a) V(s)$ 

vient (A) i'a linear constraints.

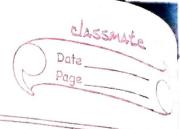
 $V(s) \geq R(s) + \gamma \leq p(s'|sa) V(s)$ 

+.a €A

Now using linear program.

s. + to V(=) > R(s) + Y \le p(0//s\_a) V(s)

YafA ses



Reporter

Support

V(s) = R(s) + Mora > P(d sa) V(s)

ath s'Es

In objective we can optimize any positive linear function of v(s) of and result above will be true

minimiz & d(s) V(s)

suf  $V(s) \ge R(s) + Y \le P(s|s,a) V(s)$   $4 af A = s \in S.$ 

d(s) is distribution over states



Adding dual variables el (2 a)

Maximize ER(s) & U(Sa). SES AGA

S+  $\leq u'(s')d(s')+Y \leq \leq p(c'|sa)$  $a \in A$   $s \in A$  u(sa)

A S (P

U (S\_a)