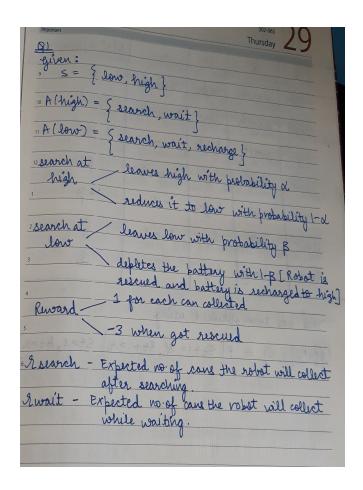
## **RL Homework 2**

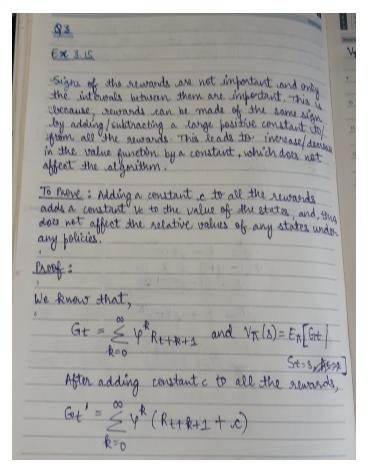
Submitted By: Snehal Gupta 2016201

Q1



whigh search low strigh wait high	rsearch rwait rsearch -3	1-A 1-B
*high search high whigh wait high wait high wait how search low whom search high	rwait	1-a
thigh search low shigh wait high low search low low	rwait	1
thigh wait high low search low high	rsearch	B
clow search low low search high		
Moro search high	-3	1-B
	rwait	1
low recharge high	0	1
This has been obtained in	sing	heama
$\frac{s}{\phi(s,a s,a)} = \rho[s+1]$	8' Rt+1	=9 St=8, At=a

According to Bellman equation for VT,
${}^{9}V_{\pi}(8) = \underbrace{\leq \pi(a s) \leq p(8',n 8,a)}_{a} \underbrace{r + \varphi V_{\pi}(8')}_{9}$ ${}^{10}V_{\pi}(8) = \underbrace{\leq \pi(a s) \leq p(8',n 8,a)}_{a} \underbrace{r + \varphi V_{\pi}(8')}_{9}$ ${}^{9}$ ${}^{11}$
$\frac{\sum \pi(a a)}{a} \frac{\sum \varphi(s',n)}{s,a} \sqrt{\pi(s')}$
$V_{\pi}(s) - \leq \pi(a s) \leq \varphi p(s', n s, a) V_{\pi}(s')$ 2  2
$= \sum_{\alpha} \sum_{\alpha} \sum_{\beta} \sum_{\beta} \sum_{\beta} \sum_{\beta} \sum_{\beta} \sum_{\beta} \sum_{\beta} \sum_{\alpha} \sum_{\beta} \sum_$
For 8=8'
scoeff of $VT(s) = (1 - \xi T(a s) \xi \psi b(s', 1 s,a)$
For 8781
Coeff of $V_{\pi}(s) = \underbrace{\{\xi \pi(a s) \xi \varphi \phi(s', \pi s, a)\}}_{s', h}$
Coeff of $V_{\pi}(s) \times V_{\pi}(s) = \underbrace{\Xi_{\pi}(a s)}_{a} \underbrace{\Xi_{\pi}(a s)}_{s',n} $



$V_{h}'(s) = E[Gt' St=s]$
$= E \left[ \sum_{k=0}^{\infty} \psi^{k} \left( R_{t+k+1} + c \right) \right] $ St=8
$= E \left\{ \begin{array}{c} & & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ \end{array} \right\} \left\{ \begin{array}{c} & & \\ & \\ & \\ \end{array} \right\} \left\{ \begin{array}{c} & & \\ & \\ & \\ \end{array} \right\} \left\{ \begin{array}{c} & \\ & \\ \end{array} \right\} \left\{ \begin{array}{c$
$= E \left\{ \begin{array}{c} \infty \\ \mathbb{R} \\ \mathbb{R} \\ \mathbb{R} \\ \mathbb{R} \end{array} \right\}$
E & V* c St=s
$= \sqrt{\pi(8)} + \leq \sqrt{k} c$
$V_{\pi}'(s) = V_{\pi}(s) + c $ [Since $0 \le y < 1$ ]
6 constant tem Ve
We can observe that the value function increases only by a constant Vc and hence, does not affect the relative values of any states under any policies.

Ex 3	
let 3	terminal time be T.
= Selve	ing the equation,
- Vils	
-	= E = VR(R++R+1+c) St=8
	= E S YR R++ R+1 St=8 +
	Keo.
.5	E & YR C St=8
	Rep
-	= Vx(s) + x(yT+1-1)
5	10 1
Edan 31	Ve
ruse, Ve	is a function of T. So, Value function for in different episodes, but same
Part eli	ede Since episodes are independent,
When the	written won't be affected.

```
\frac{\sqrt{5}}{2} \quad \sqrt{2} \quad
```