

---

```

% Achyuth Nandikotkur
% V00975928
% Question #3

clear;
clc;

alpha = 0.25;
rsearch = 5;
rwait = 1;
gamma = 0.8;
beta = 0.25;

% initial values
% vstates(1,:) holds v(low) iteration values and vstates(1, :) holds
% v(high) iteration values
vstates = [0; 0];

low = 1;
high = 2;
index = 1;

loop = 1;

while loop
    vlowwait = (rwait + (gamma* vstates(low, index)));
    vlowsearch = beta * (rsearch + (gamma * vstates(low, index))) + (1
- beta) * (-3 + (gamma * vstates(high, index)));
    vrecharge = (gamma * vstates(high, index));

    vhighwait = (rwait + (gamma * vstates(high, index)));
    vhighsearch = alpha * (rsearch + (gamma * vstates(high, index))) +
(1 - alpha) * (rsearch + (gamma * vstates(low, index)));

    % Bellman optimality equation
    vlow = max([vlowwait, vlowsearch, vrecharge]);
    vhigh = max([vhighwait, vhighsearch]);

    vstates = [vstates [vlow; vhigh]];
    index = index + 1;

    if (vstates(low, index) == vstates(low, index - 1) &&
vstates(high, index) == vstates(high, index-1))
        loop = 0;
    end
end

%action value of searching when low
qlowsearch = beta * (rsearch + (gamma * vstates(low, index))) + (1 -
beta) * (-3 + (gamma * vstates(high, index)));

%action value of waiting when low

```

---

---

```

qlowwait = (rwait + (gamma* vstates(low, index)));

%action value of recharging when low
qlowrecharge = (gamma * vstates(high, index));

%action value of searching when low
qhighsearch = alpha * (rsearch + (gamma * vstates(high, index))) + (1
- alpha) * (rsearch + (gamma * vstates(low, index)));

%action value of waiting when low
qhighwait = (rwait + (gamma * vstates(high, index)));

disp(['***** Optimal Policy *****'])
disp(['Action value of searching when low is: ',
num2str(qlowsearch)]);
disp(['Action value of waiting when low is: ', num2str(qlowwait)]);
disp(['Action value of recharging when low is: ',
num2str(qlowrecharge)]);
disp(['Action value of searching when high is: ',
num2str(qhighsearch)]);
disp(['Action value of waiting when high is: ', num2str(qhighwait)]);
fprintf('\n');
disp(['optimal value for v(low) is: ', num2str(vstates(low, index))]);
disp(['optimal value for v(high) is: ', num2str(vstates(high,
index))]);
fprintf('\n');
disp(['p(high, search) = ', num2str(1)]);
disp(['p(high, wait) = ', num2str(0)]);
disp(['p(low, search) = ', num2str(0)]);
disp(['p(low, wait) = ', num2str(0)]);
disp(['p(low, recharge) = ', num2str(1)]);

fprintf('\n')
disp(['Number of iterations ', num2str(index)])

***** Optimal Policy *****
Action value of searching when low is: 10.875
Action value of waiting when low is: 11
Action value of recharging when low is: 12.5
Action value of searching when high is: 15.625
Action value of waiting when high is: 13.5

optimal value for v(low) is: 12.5
optimal value for v(high) is: 15.625

p(high, search) = 1
p(high, wait) = 0
p(low, search) = 0
p(low, wait) = 0
p(low, recharge) = 1

Number of iterations 162

```

---

