```matlab
% Achyuth Nandikotkur
% V00975928
% Question #2

clear;
clc;

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

pihighsearch = 0.5;
pihighwait = 0.5;

pilow = {0 , 0, 0};
pilow{1} = 0.5;
pilow{2} = 0.25;
pilow{3} = 0.25;

policynew = {0, 0, 0, 0, 0};

policynew{1} = pihighsearch;
policynew{2} = pihighwait;
policynew{3} = pilow{1};
policynew{4} = pilow{2};
policynew{5} = pilow{3};


afactor = 0.001;

% 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;
outerloop = 1;
policyIndex = 1;

isPolicyStable = 0;

accuracy1 = Inf;
accuracy2 = Inf;

while outerloop
```

```matlab
    disp(['***************** Policy #', num2str(policyIndex), '
***************'])
    disp(['p(high, search) = ', num2str(pihighsearch)]);
    disp(['p(high, wait) = ', num2str(pihighwait)]);
    disp(['p(low, search) = ', num2str(pilow{1})]);
    disp(['p(low, wait) = ', num2str(pilow{2})]);
    disp(['p(low, recharge) = ', num2str(pilow{3})]);
    fprintf('\n\n')
    while loop
        vlowwait = (rwait + (gamma* vstates(low, index)));
        vlowsearch = beta * (rsearch + (gamma * vstates(low, index)))
+ (1 - beta) * (-3 + (gamma * vstates(high, index)));
        vlowrecharge = (gamma * vstates(high, index));

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

        vlow = pilow{1} * vlowsearch + pilow{2} * vlowwait + pilow{3}
* vlowrecharge;
        vhigh = pihighsearch * vhighsearch + pihighwait * vhighwait;

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

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

        disp(['value of v(low) is: ', num2str(vstates(low, index))]);
        disp(['value of v(high) is: ',num2str(vstates(high, index))]);
    end
    disp('*******************************************')

    %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)));

    if(qhighsearch == qhighwait)
        pihighsearch = 0.5;
```

```matlab
            pihighwait = 0.5;
        elseif(qhighsearch > qhighwait)
            pihighsearch = 1;
            pihighwait = 0;
        else
            pihighsearch = 0;
            pihighwait = 1;
        end

        A = [qlowsearch qlowwait qlowrecharge];
        maxval = max(A);
        lia = ismember(A,maxval);
        idx = find(lia);

        prob = 1;
        pilow = {0, 0, 0};
        for i = 1:numel(idx)
            pilow{idx} = (prob/numel(idx));
            if(size(idx) == 1)
                break;
            end
        end

        if((policynew{1} == pihighsearch && policynew{2} == pihighwait &&
 policynew{3} == pilow{1} && policynew{4} == pilow{2} && policynew{5}
 == pilow{3}))
            break;
        else
            policynew{1} = pihighsearch;
            policynew{2} = pihighwait;
            policynew{3} = pilow{1};
            policynew{4} = pilow{2};
            policynew{5} = pilow{3};
        end
        fprintf('\n');
        disp('Performing greedy improvement...')
        disp('New policy calculated, evaluating new policy..')
        fprintf('\n');

        policyIndex = policyIndex + 1;
        loop = 1;
end

fprintf('\n\n')
disp(['**************** Optimal Policy ****************'])

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

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

```
vlow = pilow{1} * vlowsearch + pilow{2} * vlowwait + pilow{3} *
 vlowrecharge;
vhigh = pihighsearch * vhighsearch + pihighwait * vhighwait;

vstates = [vstates [vlow; vhigh]];
index = index+1;
disp(['value of optimal v(low) is: ', num2str(vstates(low, index))]);
disp(['value of optimal v(high) is: ',num2str(vstates(high, index))]);

disp(['p(high, search) = ', num2str(pihighsearch)]);
disp(['p(high, wait) = ', num2str(pihighwait)]);
disp(['p(low, search) = ', num2str(pilow{1})]);
disp(['p(low, wait) = ', num2str(pilow{2})]);
disp(['p(low, recharge) = ', num2str(pilow{3})]);
disp(['**************** Optimal Policy ****************'])
```

***************** Policy #1 *****************
p(high, search) = 0.5
p(high, wait) = 0.5
p(low, search) = 0.5
p(low, wait) = 0.25
p(low, recharge) = 0.25

value of v(low) is: -0.25
value of v(high) is: 3
value of v(low) is: 1.175
value of v(high) is: 4.425
value of v(low) is: 2.315
value of v(high) is: 5.565
value of v(low) is: 3.227
value of v(high) is: 6.477
value of v(low) is: 3.9566
value of v(high) is: 7.2066
value of v(low) is: 4.5403
value of v(high) is: 7.7903
value of v(low) is: 5.0072
value of v(high) is: 8.2572
value of v(low) is: 5.3808
value of v(high) is: 8.6308
value of v(low) is: 5.6796
value of v(high) is: 8.9296
value of v(low) is: 5.9187
value of v(high) is: 9.1687
value of v(low) is: 6.11
value of v(high) is: 9.36
value of v(low) is: 6.263
value of v(high) is: 9.513
value of v(low) is: 6.3854
value of v(high) is: 9.6354
value of v(low) is: 6.4833
value of v(high) is: 9.7333
value of v(low) is: 6.5616

```
value of v(high) is: 9.8116
value of v(low) is: 6.6243
value of v(high) is: 9.8743
value of v(low) is: 6.6744
value of v(high) is: 9.9244
value of v(low) is: 6.7146
value of v(high) is: 9.9646
value of v(low) is: 6.7466
value of v(high) is: 9.9966
value of v(low) is: 6.7723
value of v(high) is: 10.0223
value of v(low) is: 6.7929
value of v(high) is: 10.0429
value of v(low) is: 6.8093
value of v(high) is: 10.0593
value of v(low) is: 6.8224
value of v(high) is: 10.0724
value of v(low) is: 6.8329
value of v(high) is: 10.0829
value of v(low) is: 6.8414
value of v(high) is: 10.0914
value of v(low) is: 6.8481
value of v(high) is: 10.0981
value of v(low) is: 6.8535
value of v(high) is: 10.1035
value of v(low) is: 6.8578
value of v(high) is: 10.1078
value of v(low) is: 6.8612
value of v(high) is: 10.1112
value of v(low) is: 6.864
value of v(high) is: 10.114
value of v(low) is: 6.8662
value of v(high) is: 10.1162
value of v(low) is: 6.8679
value of v(high) is: 10.1179
value of v(low) is: 6.8694
value of v(high) is: 10.1194
value of v(low) is: 6.8705
value of v(high) is: 10.1205
value of v(low) is: 6.8714
value of v(high) is: 10.1214
*******************************************

Performing greedy improvement...
New policy calculated, evaluating new policy..

***************** Policy #2 ****************
p(high, search) = 1
p(high, wait) = 0
p(low, search) = 0
p(low, wait) = 0
p(low, recharge) = 1
```

```
value of v(low) is: 8.0971
value of v(high) is: 11.1471
value of v(low) is: 8.9177
value of v(high) is: 12.0877
value of v(low) is: 9.6702
value of v(high) is: 12.7682
value of v(low) is: 10.2145
value of v(high) is: 13.3557
value of v(low) is: 10.6846
value of v(high) is: 13.7999
value of v(low) is: 11.0399
value of v(high) is: 14.1707
value of v(low) is: 11.3366
value of v(high) is: 14.4581
value of v(low) is: 11.5665
value of v(high) is: 14.6936
value of v(low) is: 11.7548
value of v(high) is: 14.8786
value of v(low) is: 11.9029
value of v(high) is: 15.0286
value of v(low) is: 12.0229
value of v(high) is: 15.1474
value of v(low) is: 12.118
value of v(high) is: 15.2432
value of v(low) is: 12.1946
value of v(high) is: 15.3194
value of v(low) is: 12.2555
value of v(high) is: 15.3806
value of v(low) is: 12.3045
value of v(high) is: 15.4294
value of v(low) is: 12.3436
value of v(high) is: 15.4686
value of v(low) is: 12.3749
value of v(high) is: 15.4999
value of v(low) is: 12.3999
value of v(high) is: 15.5249
value of v(low) is: 12.4199
value of v(high) is: 15.5449
value of v(low) is: 12.4359
value of v(high) is: 15.5609
value of v(low) is: 12.4487
value of v(high) is: 15.5737
value of v(low) is: 12.459
value of v(high) is: 15.584
value of v(low) is: 12.4672
value of v(high) is: 15.5922
value of v(low) is: 12.4738
value of v(high) is: 15.5988
value of v(low) is: 12.479
value of v(high) is: 15.604
value of v(low) is: 12.4832
value of v(high) is: 15.6082
value of v(low) is: 12.4866
value of v(high) is: 15.6116
```

```
value of v(low) is: 12.4893
value of v(high) is: 15.6143
value of v(low) is: 12.4914
value of v(high) is: 15.6164
value of v(low) is: 12.4931
value of v(high) is: 15.6181
value of v(low) is: 12.4945
value of v(high) is: 15.6195
value of v(low) is: 12.4956
value of v(high) is: 15.6206
value of v(low) is: 12.4965
value of v(high) is: 15.6215
*********************************************


***************** Optimal Policy *****************
value of optimal v(low) is: 12.4972
value of optimal v(high) is: 15.6222
p(high, search) = 1
p(high, wait) = 0
p(low, search) = 0
p(low, wait) = 0
p(low, recharge) = 1
***************** Optimal Policy *****************
```

Published with MATLAB® R2021a