The nonzero value function is

3	$2.032705\,\mathrm{e}{+}01$
11	$2.148484\,\mathrm{e}{+}01$
12	$2.089052\,\mathrm{e}{+}01$
15	$2.691525\,\mathrm{e}{+}01$
16	$2.768148\,\mathrm{e}{+}01$
17	$2.846952\,\mathrm{e}{+}01$
20	$2.209646\,\mathrm{e}{+}01$
22	$2.474148\,\mathrm{e}{+}01$
23	$2.544584\,\mathrm{e}{+}01$
24	$2.617024\,\mathrm{e}{+}01$
26	$2.928000\,\mathrm{e}{+}01$
29	$2.272550\,\mathrm{e}{+}01$
30	$2.337246\,\mathrm{e}{+}01$
31	$2.405617\,\mathrm{e}{+}01$
34	$3.097082\,\mathrm{e}{+}01$
35	$3.011354\,\mathrm{e}{+}01$
39	$2.269388\,\mathrm{e}{+}01$
43	$3.185250\,\mathrm{e}{+}01$
47	$-4.000000\,\mathrm{e}{+}01$
48	$2.086217\,\mathrm{e}{+}01$
49	$-4.000000\,\mathrm{e}{+}01$
51	$-4.000000\,\mathrm{e}{+}01$
52	$3.275928\,\mathrm{e}{+}01$
53	$3.388155\,\mathrm{e}{+}01$
56	$2.081641\,\mathrm{e}{+}01$
57	$2.468081\,\mathrm{e}{+}01$
58	$2.567706\,\mathrm{e}{+}01$
59	$2.996586\,\mathrm{e}{+}01$
60	$3.091243\,\mathrm{e}{+}01$
61	$3.562801\,\mathrm{e}{+}01$
62	$3.485108\,\mathrm{e}{+}01$
65	$-4.000000\,\mathrm{e}{+}01$
66	$2.081641\,\mathrm{e}{+}01$
67	$-4.000000\,\mathrm{e}{+}01$
69	$-4.000000\,\mathrm{e}{+}01$
70	$3.684160\,\mathrm{e}{+}01$
71	$3.582236\,\mathrm{e}{+}01$
79	$4.000000\,\mathrm{e}{+}01$

The optimal policy is

3	3
11	3
12	2
15	4
16	4
17	3
20	3
22	4

```
23      4
24      1
26      3
29      4
30      4
31      1
34      3
35      2
39      1
43      3
47      1
48      3
49      1
51      1
52      3
53      3
56      4
57      4
58      4
59      4
60      4
61      3
62      3
65      1
66      1
67      1
69      1
70      3
71      2
79      1
```

The following is the Matlab source code
————————l——-oad the data————————

```
function [Pa1,Pa2,Pa3,Pa4,Rewards] = cse_hw6_load()
% function [Pa1,Pa2,Pa3,Pa4] = cse_hw6_load()
% load the transition matrices, pai corresponding to the transition matrix
% of action ai

S=81;

Rewards = importdata('rewards.txt');

A1 = importdata('prob_a1.txt');
A2 = importdata('prob_a2.txt');
A3 = importdata('prob_a3.txt');
A4 = importdata('prob_a4.txt');

spA1 = spconvert(A1);
spA2 = spconvert(A2);
```

```
spA3 = spconvert(A3);
spA4 = spconvert(A4);

Pa1 = full(spA1);
Pa2 = full(spA2);
Pa3 = full(spA3);
Pa4 = full(spA4);
```

————————calculate value function and optimal policy—————-

```
%cse_hw6_a.m

iter = 10000;%number of iterations
state = 81;%number of states
V = zeros(state,1);%initialize value function
gama = 0.975;

[Pa1,Pa2,Pa3,Pa4,Rewards] = cse_hw6_load();%load the transition matrix
for i = 1:iter
    V = Rewards + gama*max([Pa1*V Pa2*V Pa3*V Pa4*V],[],2);
end
FID1 = fopen('valuefunction','w+');
FID2 = fopen('nonzerovaluefunction','w+');
FID3 = fopen('optimalpolicy','w+');
for i=1:state
        fprintf(FID1, '%d %d \n', i, V(i));
end
fclose(FID1);

for i=1:state
    if(V(i) ~= 0)
        fprintf(FID2, '%-5d %-5d \n', i, V(i));
    end
end
fclose(FID2);

[Y, PIE] = max([Pa1*V Pa2*V Pa3*V Pa4*V],[],2);
for i=1:state
    if(V(i) ~= 0)
        fprintf(FID3, '%-5d %-5d \n', i, PIE(i));
    end
end
fclose(FID3);
```