# Exercises - Chapter 4

Carl Fredriksson, c@msp.se

## Exercise 4.1

In Example 4.1, if $\pi$ is the equiprobable random policy, what is $q_\pi(11, down)$? What is $q_\pi(7, down)$?

**My answer:**

$$\begin{aligned} q_\pi(11, down) &= \sum_{s',r} p(s', r|11, down)\big[r + v_\pi(s')\big] \\ &= -1 + v_\pi(terminal) \\ &= -1 \end{aligned}$$

$$\begin{aligned} q_\pi(7, down) &= \sum_{s',r} p(s', r|7, down)\big[r + v_\pi(s')\big] \\ &= -1 + v_\pi(11) \\ &= -15 \end{aligned}$$

The values for $v_\pi$ can be seen in figure 4.1.

## Exercise 4.2

In Example 4.1, suppose a new state 15 is added to the gridworld just below state 13, and its actions, left, up, right, and down, take the agent to states 12, 13, 14, and 15, respectively. Assume that the transitions from the original states are unchanged. What, then, is $v_\pi(15)$ for the equiprobable random policy? Now suppose the dynamics of state 13 are also changed, such that action down from state 13 takes the agent to the new state 15. What is $v_\pi(15)$ for the equiprobable random policy in this case?

**My answer:**

For the case that the transitions from the original states are unchanged:

$$\begin{aligned} v_\pi(15) &= \sum_a \pi(a|15) \sum_{s',r} p(s', r|15, a)\big[r + v_\pi(s')\big] \\ &= 0.25\big[-1 + v_\pi(12)\big] + 0.25\big[-1 + v_\pi(13)\big] + 0.25\big[-1 + v_\pi(14)\big] + 0.25\big[-1 + v_\pi(15)\big] \\ &= -1 + 0.25\big[-22 - 20 - 14 + v_\pi(15)\big] \\ &= -15 + 0.25 v_\pi(15) \\ &= \frac{-15}{0.75} \\ &= -20 \end{aligned}$$

For the case where the transitions from state 13 are changed:

I wrote a program implementing iterative policy evaluation for this example. It got the same results as in the book for the original example and validated $v_\pi(15) = -20$ for the case that transitions from the original states are unchanged. When running the program for the case where transitions for state 13 are changed I found that the state values didn't change. This makes sense, since the state values for the next state after selecting an action is identical for state 13 and state 15. Thus we still have:

$$v_\pi(15) = -20$$

## Exercise 4.3

What are the equations analogous to (4.3), (4.4), and (4.5), but for action-value functions instead of state-value functions?

**My answer:**

$$\begin{aligned} q_\pi(s,a) &= \mathbb{E}_\pi[G_t|S_t=s,A_t=a] \\ &= \mathbb{E}_\pi[R_t+\gamma G_{t+1}|S_t=s,A_t=a] \\ &= \mathbb{E}[R_t+\gamma v_\pi(S_{t+1})|S_t=s,A_t=a] \\ &= \mathbb{E}[R_t+\gamma \sum_{a'\in\mathcal{A}(S_{t+1})} \pi(a'|S_{t+1})q_\pi(S_{t+1},a')|S_t=s,A_t=a] \\ &= \sum_{s',r} p(s',r|s,a)\big[r+\gamma\sum_{a'}\pi(a'|s')q_\pi(s',a')\big] \end{aligned}$$

$$\begin{aligned} q_{k+1}(s,a) &= \mathbb{E}[R_t+\gamma \sum_{a'\in\mathcal{A}(S_{t+1})} \pi(a'|S_{t+1})q_k(S_{t+1},a')|S_t=s,A_t=a] \\ &= \sum_{s',r} p(s',r|s,a)\big[r+\gamma\sum_{a'}\pi(a'|s')q_k(s',a')\big] \end{aligned}$$

## Exercise 4.4

The policy iteration algorithm on page 80 has a subtle bug in that it may never terminate if the policy continually switches between two or more policies that are equally good. This is okay for pedagogy, but not for actual use. Modify the pseudocode so that convergence is guaranteed.

**My answer:**

```
1. Initialization
    ○ $V(s) \in \mathbb{R}$ and $\pi(s) \in \mathcal{A}(s)$ arbitrarily for all $s \in \mathcal{S}$; $V(terminal) = 0$
2. Policy Evaluation
    ○ Loop:
        ▪ $\Delta \leftarrow 0$
        ▪ For each $s \in \mathcal{S}$:
            ▪ $v \leftarrow V(s)$
            ▪ $V(s) \leftarrow \sum_{s',r} p(s',r|s,\pi(s))\big[r+\gamma V(s')\big]$
            ▪ $\Delta \leftarrow \max(\Delta, |v - V(s)|)$
    ○ until $\Delta < \theta$ (a small positive number determining the accuracy of estimation)
3. Policy Improvement
    ○ $policy\text{-}stable \leftarrow true$
    ○ For each $s \in \mathcal{S}$:
        ▪ $\pi(s) \leftarrow \arg\max_a \sum_{s',r} p(s',r|s,a)\big[r+\gamma V(s')\big]$
        ▪ $Q(s,\pi(s)) \leftarrow \sum_{s',r} p(s',r|s,\pi(s))\big[r+\gamma V(s')\big]$
        ▪ If $Q(s,\pi(s)) > V(s) + \theta$, then $policy\text{-}stable \leftarrow false$
    ○ If $policy\text{-}stable$, then stop and return $V \approx v_*$ and $\pi \approx \pi_*$; else go to 2
```

## Exercise 4.5

How would policy iteration be defined for action values? Give a complete algorithm for computing $q_*$, analogous to that on page 80 for computing $v_*$. Please pay special attention to this exercise, because the ideas involved will be used throughout the rest of the book.

**My answer:**

```
1. Initialization
    ○ $Q(s,a) \in \mathbb{R}$ arbitrarily for all $s \in \mathcal{S}, a \in \mathcal{A}(s)$
    ○ $\pi(s) \in \mathcal{A}(s)$ arbitrarily for all $s \in \mathcal{S}$
    ○ $Q(terminal, \_) = 0$
2. Policy Evaluation
    ○ Loop:
        ▪ $\Delta \leftarrow 0$
        ▪ For each $s \in \mathcal{S}, a \in \mathcal{A}(s)$:
            ▪ $q \leftarrow Q(s,a)$
            ▪ $Q(s,a) \leftarrow \sum_{s',r} p(s',r|s,a)\big[r + \gamma Q(s', \pi(s'))\big]$
            ▪ $\Delta \leftarrow \max(\Delta, |q - Q(s,a)|)$
    ○ until $\Delta < \theta$ (a small positive number determining the accuracy of estimation)
3. Policy Improvement
    ○ $policy\text{-}stable \leftarrow true$
    ○ For each $s \in \mathcal{S}$:
        ▪ $old\text{-}action \leftarrow \pi(s)$
        ▪ $\pi(s) \leftarrow \arg\max_a Q(s,a)$
        ▪ If $Q(s, \pi(s)) > Q(s, old\text{-}action) + \theta$, then $policy\text{-}stable \leftarrow false$
    ○ If $policy\text{-}stable$, then stop and return $Q \approx q_*$ and $\pi \approx \pi_*$; else go to 2
```

## Exercise 4.6

Suppose you are restricted to considering only policies that are $\epsilon - soft$, meaning that the probability of selecting each action in each state, s, is at least $\epsilon/|A(s)|$. Describe qualitatively the changes that would be required in each of the steps 3, 2, and 1, in that order, of the policy iteration algorithm for $v_*$ on page 80.

**My answer:**

In step 3: rather than setting a deterministic policy selecting the action $a$ that maximizes $\sum_{s',r} p(s',r|s,a)\big[r + \gamma V(s')\big]$ in each state, we would would give the maximum possible probability to the maximizing action, and the other actions would get the least possible probability $\epsilon/|A(s)|$.

In step 2: we would have to take into account that the policy is not deterministic and change:

$$V(s) \leftarrow \sum_{s',r} p(s',r|s,\pi(s))\big[r + \gamma V(s')\big]$$

to:

$$V(s) \leftarrow \sum_{a \in \mathcal{A}(s)} \pi(a|s) \sum_{s',r} p(s',r|s,a)\big[r + \gamma V(s')\big]$$

In step 1: rather than initializing an arbitrary deterministic policy, we would have to initialize $\pi(a|s)$ so that all actions in every state has at least the minimum probability $\epsilon/|A(s)|$ of being selected.

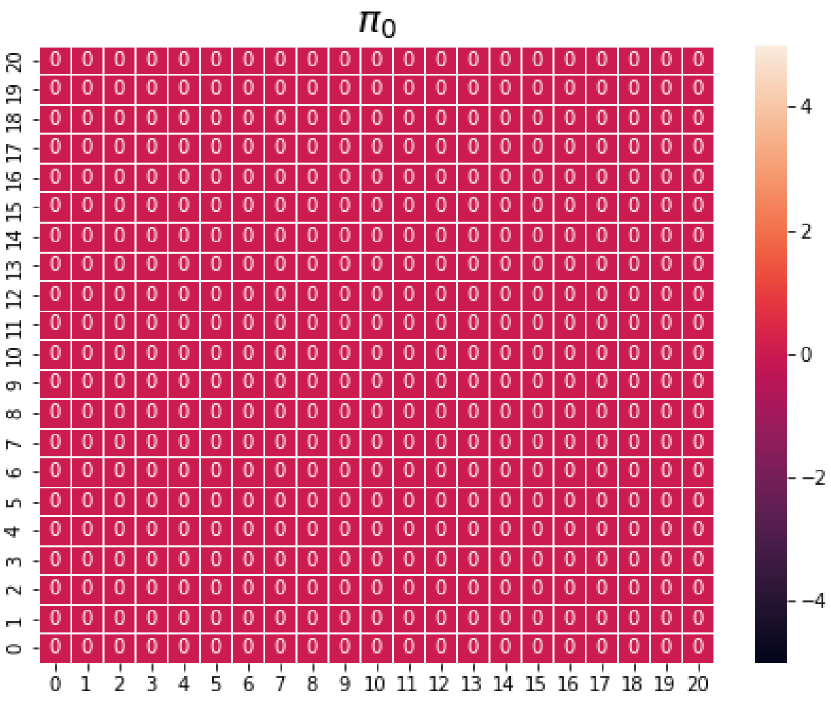
## Exercise 4.7 (programming)

Write a program for policy iteration and re-solve Jack's car rental problem with the following changes. One of Jack's employees at the first location rides a bus home each night and lives near the second location. She is happy to shuttle one car to the second location for free. Each additional car still costs $2, as do all cars moved in the other direction. In addition, Jack has limited parking space at each location. If more than 10 cars are kept overnight at a location (after any moving of cars), then an additional cost of $4 must be incurred to use a second parking lot (independent of how many cars are kept there). These sorts of nonlinearities and arbitrary dynamics often occur in real problems and cannot easily be handled by optimization methods other than dynamic programming. To check your program, first replicate the results given for the original problem.

**My answer:**

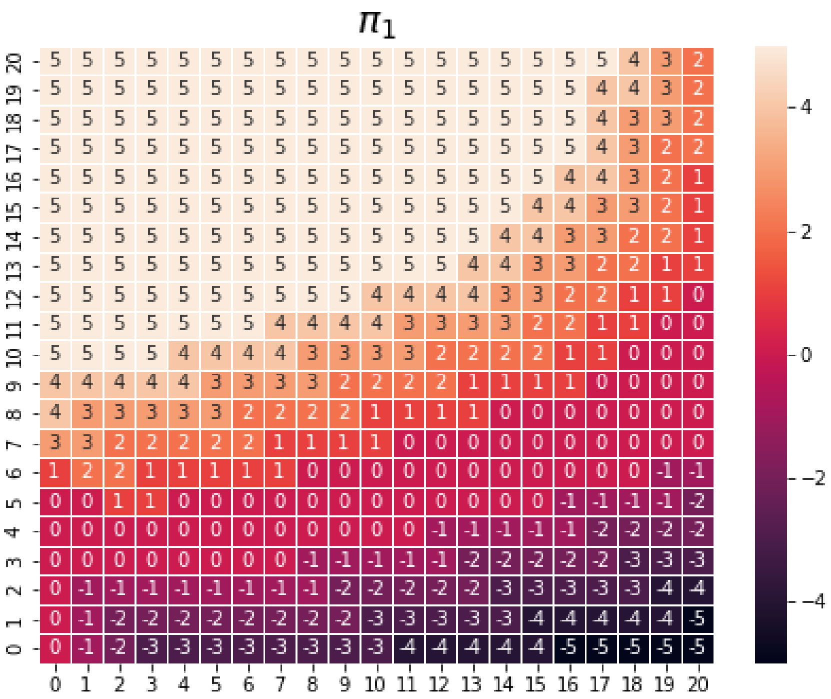
Original problem:

I got similar but not exactly the same results as in the book. $\pi_4$ is stable.



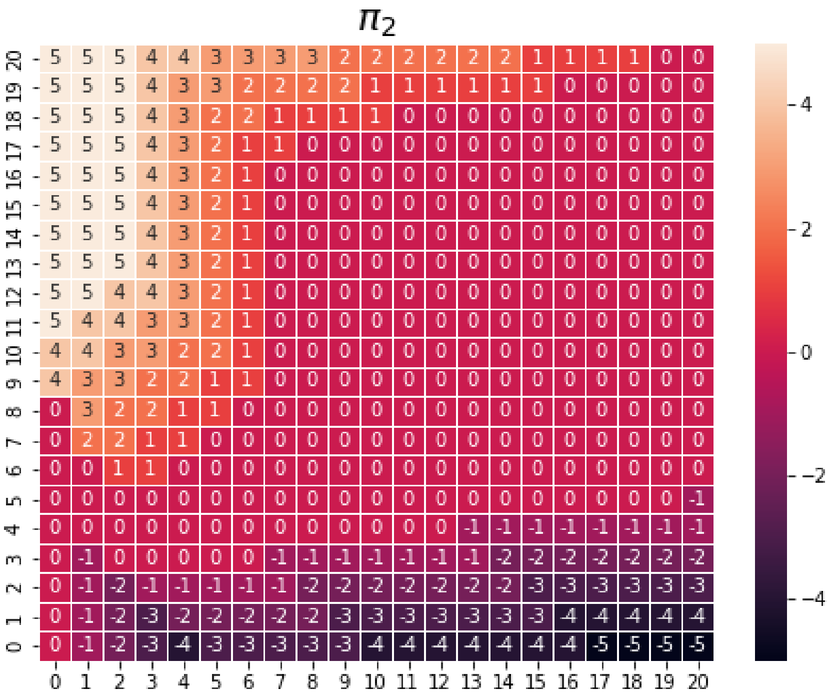
$\pi_0$

|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 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 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

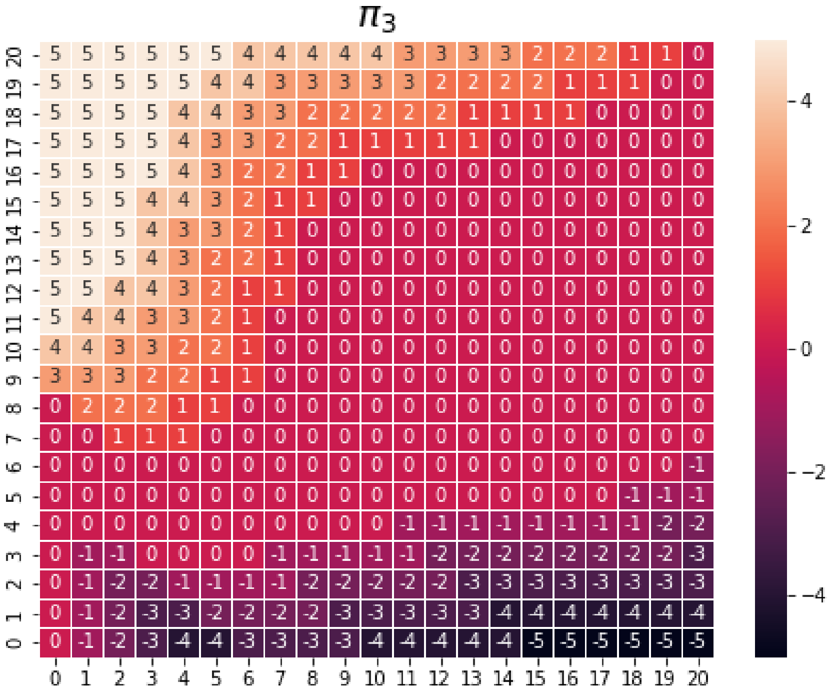


$\pi_1$

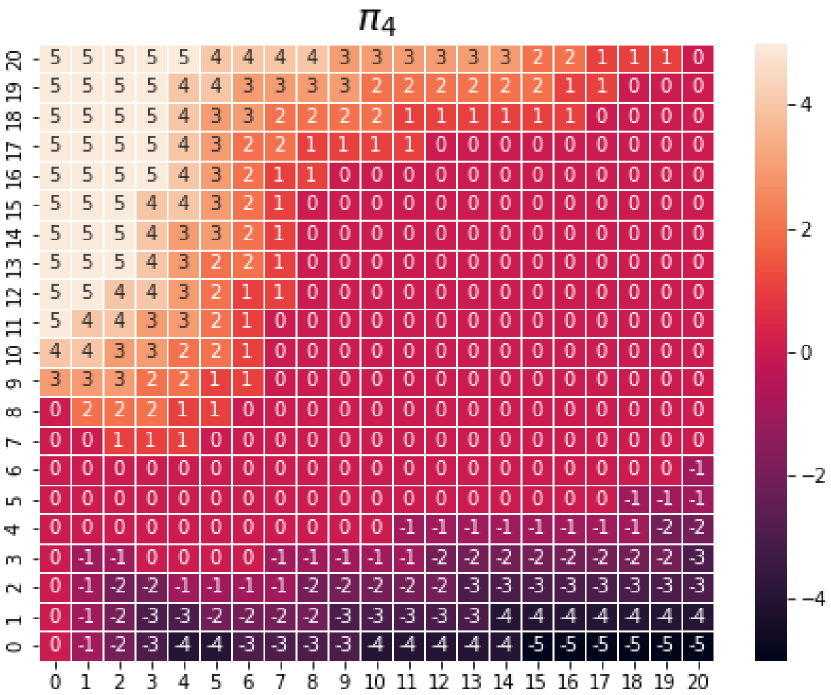
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 2 |
| 19 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 2 |
| 18 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 3 | 2 |
| 17 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 2 |
| 16 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 |
| 15 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 2 | 1 |
| 14 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 2 | 2 | 1 |
| 13 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 2 | 2 | 1 | 1 |
| 12 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 0 |
| 11 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 0 | 0 |
| 10 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 0 | 0 | 0 |
| 9 | 4 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 8 | 4 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 |
| 5 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -2 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 |
| 2 | 0 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -4 | -4 |
| 1 | 0 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -5 |
| 0 | 0 | -1 | -2 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 |


$\pi_2$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 |
| 19 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 18 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 5 | 4 | 4 | 3 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 4 | 4 | 3 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 3 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| 3 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 |
| 2 | 0 | -1 | -2 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -3 |
| 1 | 0 | -1 | -2 | -3 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 |
| 0 | 0 | -1 | -2 | -3 | -4 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 |


$\pi_3$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 1 | 1 | 0 |
| 19 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 0 | 0 |
| 18 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 17 | 5 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 5 | 4 | 4 | 3 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 4 | 4 | 3 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 2 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -2 | -2 |
| 3 | 0 | -1 | -1 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -3 |
| 2 | 0 | -1 | -2 | -2 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 |
| 1 | 0 | -1 | -2 | -3 | -3 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -4 | -4 |
| 0 | 0 | -1 | -2 | -3 | -4 | -4 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 |



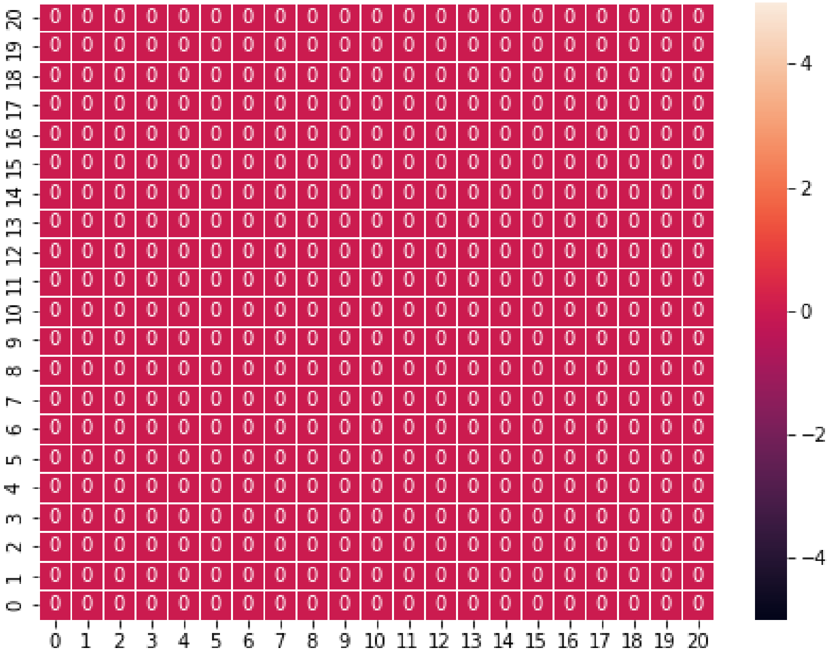
$\pi_4$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 0 |
| 19 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 0 | 0 | 0 |
| 18 | 5 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 17 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 5 | 4 | 4 | 3 | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 4 | 4 | 3 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 2 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -2 | -2 |
| 3 | 0 | -1 | -1 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -3 |
| 2 | 0 | -1 | -2 | -2 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 |
| 1 | 0 | -1 | -2 | -3 | -3 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -4 | -4 |
| 0 | 0 | -1 | -2 | -3 | -4 | -4 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 |

Updated problem:

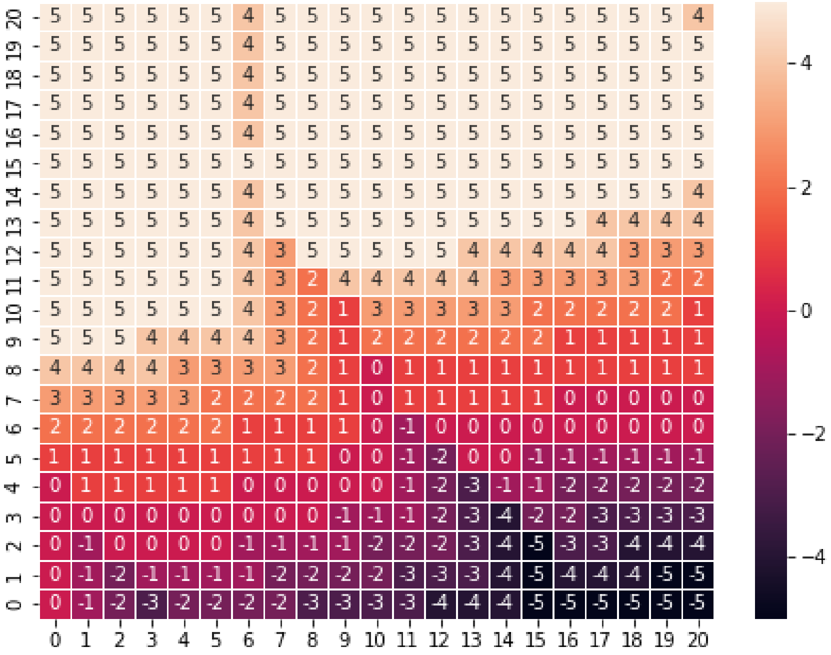
$\pi_5$ is stable.

$\pi_0$



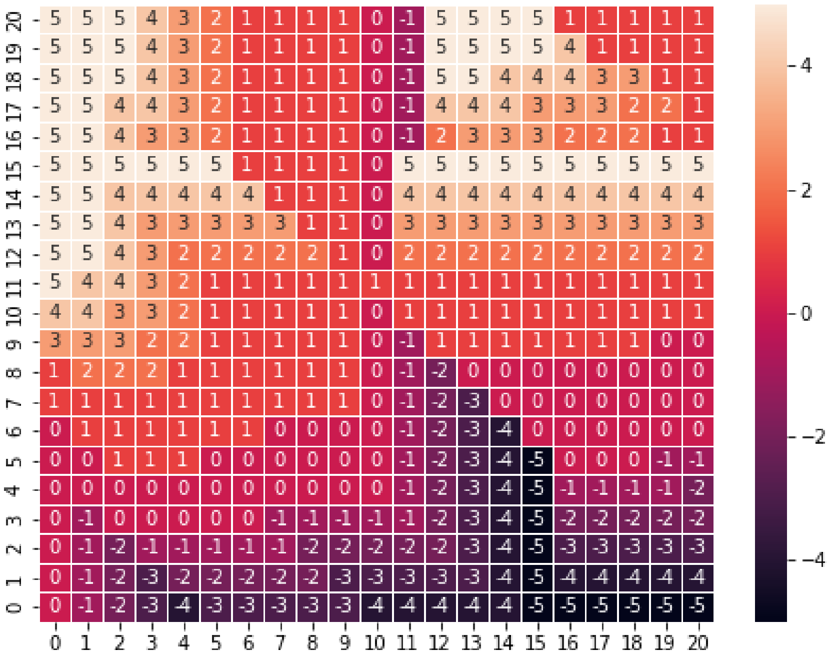
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 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 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$\pi_1$



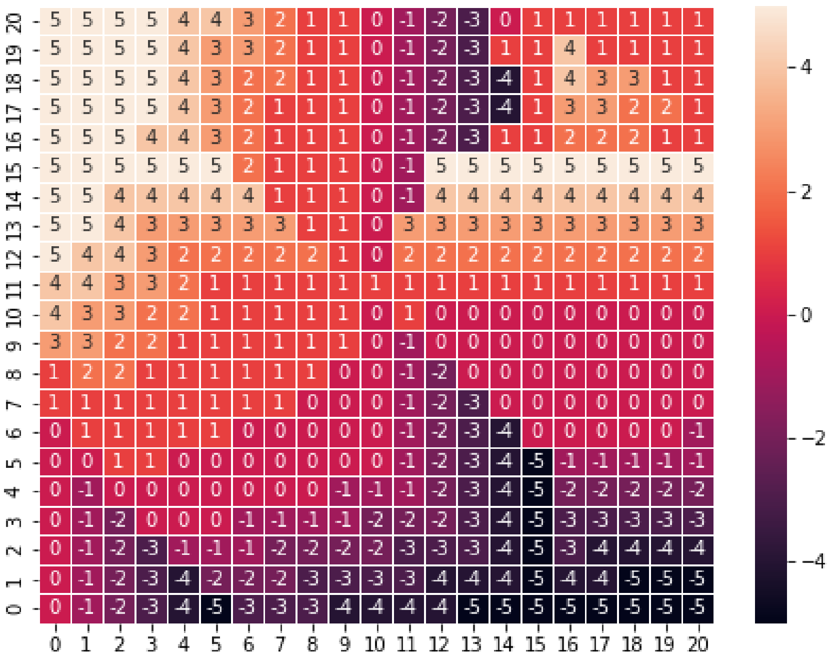
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 |
| 19 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 18 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 17 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 16 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 15 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 14 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 |
| 13 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 |
| 12 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 3 | 3 | 3 |
| 11 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 4 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 2 | 2 |
| 10 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 1 |
| 9 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 3 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| 8 | 4 | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 6 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | -1 | -2 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -1 |
| 4 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -1 | -1 | -2 | -2 | -2 | -2 | -2 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -2 | -3 | -4 | -2 | -2 | -3 | -3 | -3 | -3 |
| 2 | 0 | -1 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -3 | -4 | -5 | -3 | -3 | -4 | -4 | -4 |
| 1 | 0 | -1 | -2 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -4 | -5 | -4 | -4 | -4 | -5 | -5 |
| 0 | 0 | -1 | -2 | -3 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 |

$\pi_2$

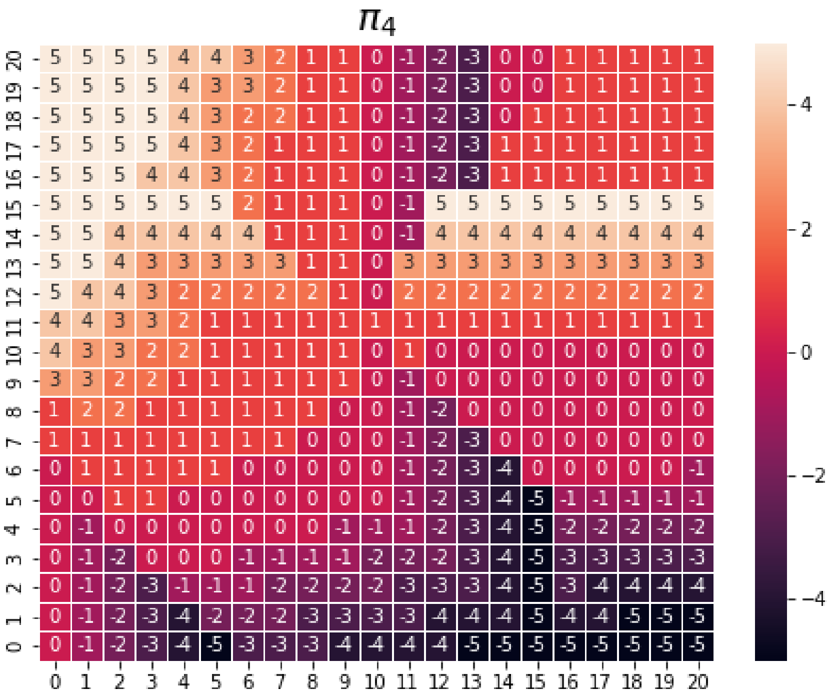


| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | 0 | -1 | 5 | 5 | 5 | 5 | 1 | 1 | 1 | 1 | 1 |
| 19 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | 0 | -1 | 5 | 5 | 5 | 5 | 4 | 1 | 1 | 1 | 1 |
| 18 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | 0 | -1 | 5 | 5 | 4 | 4 | 4 | 3 | 3 | 1 | 1 |
| 17 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | 0 | -1 | 4 | 4 | 4 | 3 | 3 | 3 | 2 | 2 | 1 |
| 16 | 5 | 5 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | 0 | -1 | 2 | 3 | 3 | 3 | 2 | 2 | 2 | 1 | 1 |
| 15 | 5 | 5 | 5 | 5 | 5 | 5 | 1 | 1 | 1 | 1 | 0 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 14 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 1 | 1 | 1 | 0 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 13 | 5 | 5 | 4 | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 0 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 12 | 5 | 5 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 11 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 4 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 3 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | -1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 8 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | -1 | -2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | -5 | 0 | 0 | 0 | -1 | -1 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | -5 | -1 | -1 | -1 | -1 | -2 |
| 3 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 | -2 | -3 | -4 | -5 | -2 | -2 | -2 | -2 | -2 |
| 2 | 0 | -1 | -2 | -1 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -3 | -3 |
| 1 | 0 | -1 | -2 | -3 | -2 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -3 | -4 | -5 | -4 | -4 | -4 | -4 | -4 |
| 0 | 0 | -1 | -2 | -3 | -4 | -3 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 |

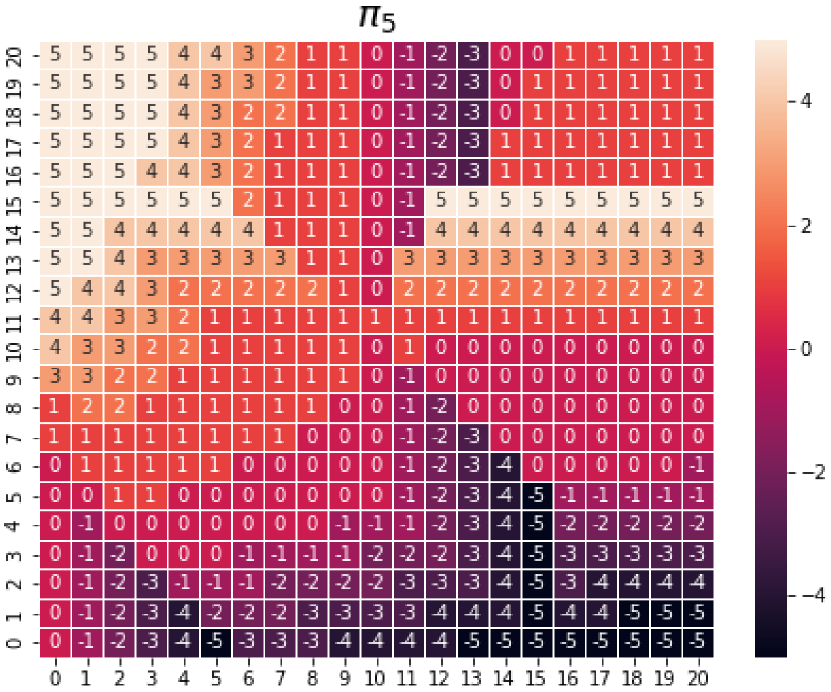
$\pi_3$



| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 19 | 5 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 1 | 1 | 4 | 1 | 1 | 1 | 1 |
| 18 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | -4 | 1 | 4 | 3 | 3 | 1 | 1 |
| 17 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | -4 | 1 | 3 | 3 | 2 | 2 | 1 |
| 16 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | 1 | 1 | 2 | 2 | 2 | 1 | 1 |
| 15 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 1 | 1 | 1 | 0 | -1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 14 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 1 | 1 | 1 | 0 | -1 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 13 | 5 | 5 | 4 | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 0 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 12 | 5 | 4 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 11 | 4 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | -1 | -2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | -1 | -2 | -3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | 0 | 0 | 0 | 0 | 0 | -1 |
| 5 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | -5 | -1 | -1 | -1 | -1 | -1 |
| 4 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -2 | -3 | -4 | -5 | -2 | -2 | -2 | -2 | -2 |
| 3 | 0 | -1 | -2 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -3 | -3 |
| 2 | 0 | -1 | -2 | -3 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -4 | -5 | -3 | -4 | -4 | -4 | -4 |
| 1 | 0 | -1 | -2 | -3 | -4 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -5 | -4 | -4 | -5 | -5 | -5 |
| 0 | 0 | -1 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 |



| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 19 | 5 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 18 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 17 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 16 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 15 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 1 | 1 | 1 | 0 | -1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 14 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 1 | 1 | 1 | 0 | -1 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 13 | 5 | 5 | 4 | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 0 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 12 | 5 | 4 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 11 | 4 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | -1 | -2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | -1 | -2 | -3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | 0 | 0 | 0 | 0 | 0 | -1 |
| 5 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | -5 | -1 | -1 | -1 | -1 | -1 |
| 4 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -2 | -3 | -4 | -5 | -2 | -2 | -2 | -2 | -2 |
| 3 | 0 | -1 | -2 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -3 | -3 |
| 2 | 0 | -1 | -2 | -3 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -4 | -5 | -3 | -4 | -4 | -4 | -4 |
| 1 | 0 | -1 | -2 | -3 | -4 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -5 | -4 | -4 | -5 | -5 | -5 |
| 0 | 0 | -1 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 |



| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 19 | 5 | 5 | 5 | 5 | 4 | 3 | 3 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 18 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 2 | 1 | 1 | 0 | -1 | -2 | -3 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 17 | 5 | 5 | 5 | 5 | 4 | 3 | 2 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 16 | 5 | 5 | 5 | 4 | 4 | 3 | 2 | 1 | 1 | 1 | 0 | -1 | -2 | -3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 15 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 1 | 1 | 1 | 0 | -1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 14 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 1 | 1 | 1 | 0 | -1 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 13 | 5 | 5 | 4 | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 0 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 12 | 5 | 4 | 4 | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 11 | 4 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 4 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | -1 | -2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | -1 | -2 | -3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | 0 | 0 | 0 | 0 | 0 | -1 |
| 5 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -2 | -3 | -4 | -5 | -1 | -1 | -1 | -1 | -1 |
| 4 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -2 | -3 | -4 | -5 | -2 | -2 | -2 | -2 | -2 |
| 3 | 0 | -1 | -2 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -2 | -2 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -3 | -3 |
| 2 | 0 | -1 | -2 | -3 | -1 | -1 | -1 | -2 | -2 | -2 | -2 | -3 | -3 | -3 | -4 | -5 | -3 | -4 | -4 | -4 | -4 |
| 1 | 0 | -1 | -2 | -3 | -4 | -2 | -2 | -2 | -3 | -3 | -3 | -3 | -4 | -4 | -4 | -5 | -4 | -4 | -5 | -5 | -5 |
| 0 | 0 | -1 | -2 | -3 | -4 | -5 | -3 | -3 | -3 | -4 | -4 | -4 | -4 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 |

## Exercise 4.8

Why does the optimal policy for the gambler's problem have such a curious form? In particular, for capital of 50 it bets it all on one flip, but for capital of 51 it does not. Why is this a good policy?

**My answer:**

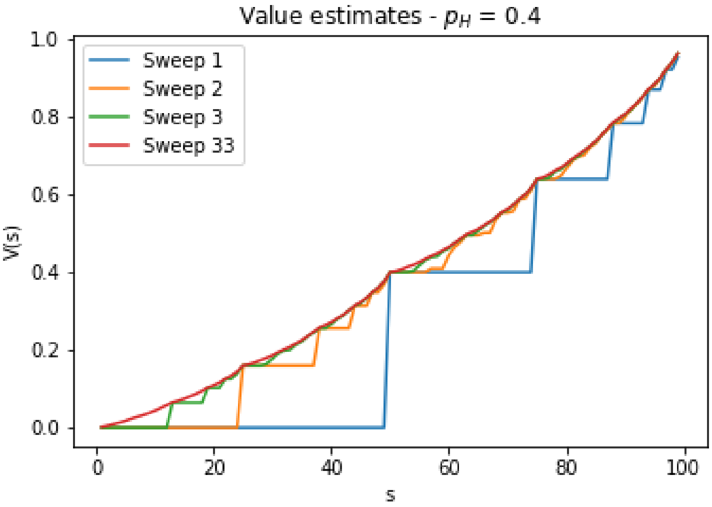
With 50 capital you have exactly enough to make a bet such that if the coin comes up heads you win. There is no way to split it up and bet multiple times to increase the probability of winning ($p_h$). With 51 you have enough to make a bet that could immediately win. However, you also have to opportunity of using the 1 extra capital to try to build up towards 75, where you could make two flips in a row that both could make you win. 25 first and then 50 if the first one fails. On the other hand, if you immediately bet 50 and lose, you can try to build up the 1 left over to 25. Building up from 1 to 25 is equally unlikely as from 51 to 75, but the former lets you make a 25 bet to get to 50 while the latter lets you make a 25 bet to win, the latter is clearly preferable.

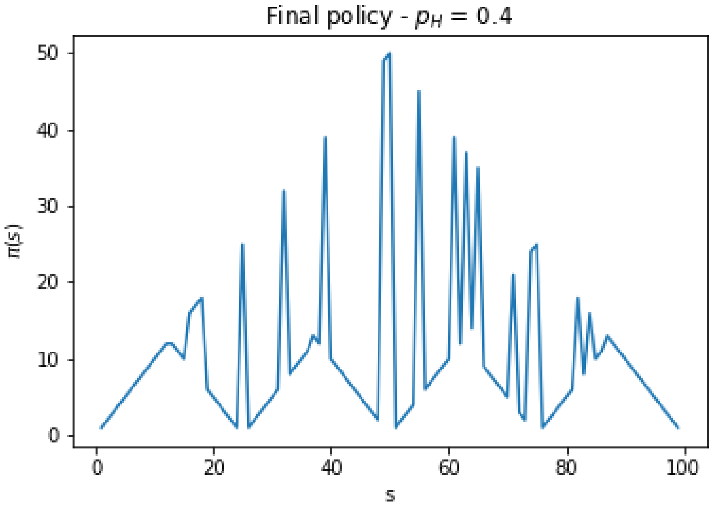
## Exercise 4.9 (programming)

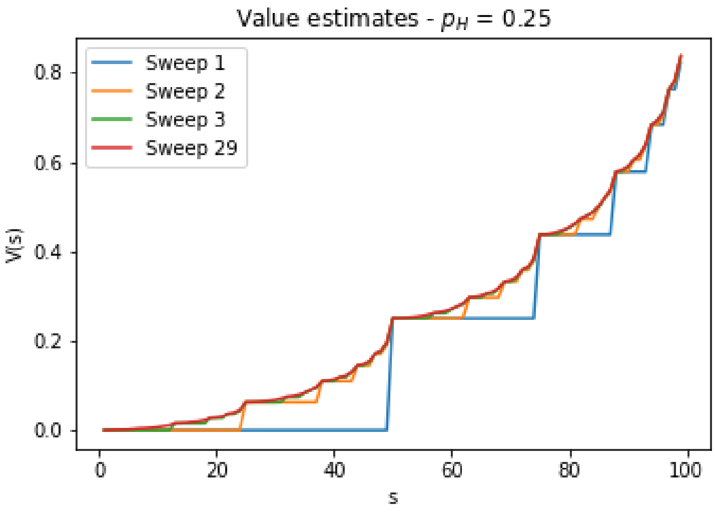
Implement value iteration for the gambler's problem and solve it for $p_h = 0.25$ and $p_h = 0.55$. In programming, you may find it convenient to introduce two dummy states corresponding to termination with capital of 0 and 100, giving them values of 0 and 1 respectively. Show your results graphically, as in Figure 4.3. Are your results stable as $\theta \to 0$?

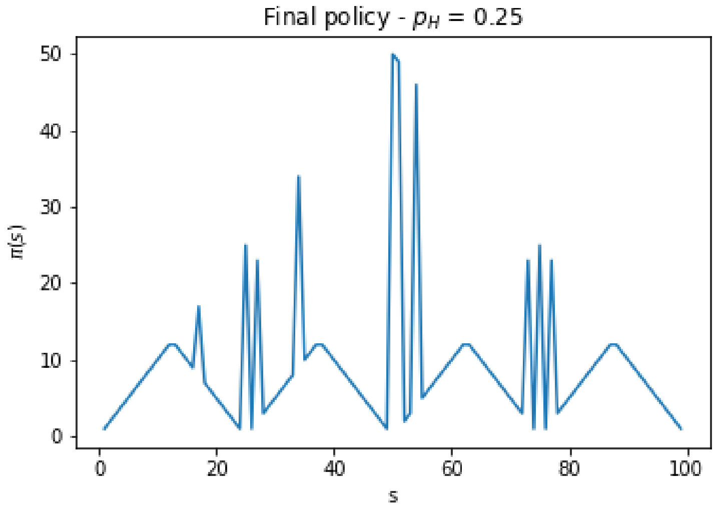
**My answer:**

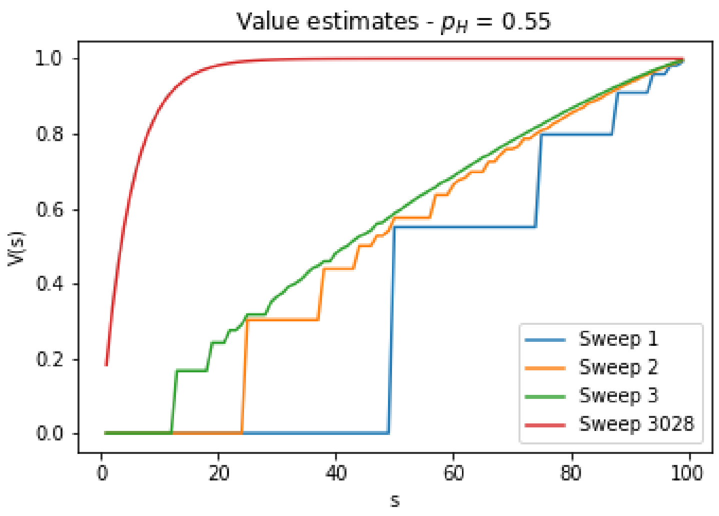
Final policies for $p_h = 0.4$ and $p_h = 0.25$ have similar shapes with $\pi(25) = 25$, $\pi(50) = 50$, $\pi(75) = 25$. However, when looking at the final $V(s)$ it seems like estimates has prematurely converged due to precision limits in the data type (np.float64). The final policy for $p_h = 0.55$ is completely different, with $\pi(s) = 1$ for all $s$. The key difference is whether $p_h > 0.5$.

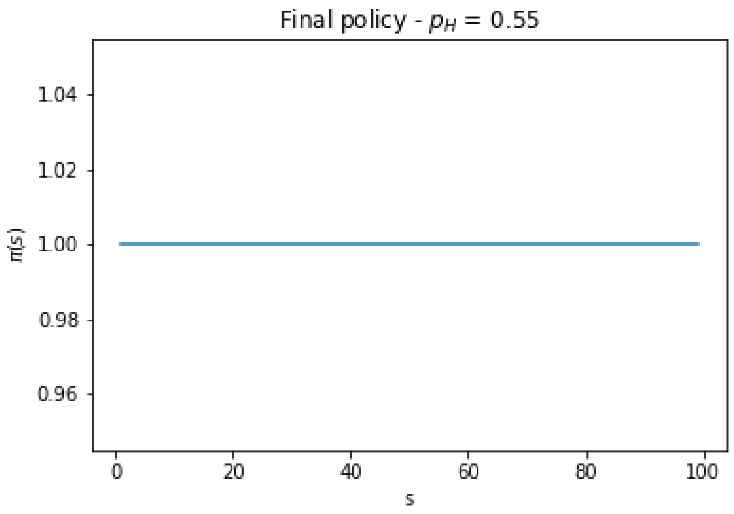





Value estimates - $p_H$ = 0.25
Sweep 1
Sweep 2
Sweep 3
Sweep 29
V(s)
0.0
0.2
0.4
0.6
0.8
0
20
40
60
80
100
s


Final policy - $p_H$ = 0.25
$\pi(s)$
0
10
20
30
40
50
0
20
40
60
80
100
s





## Exercise 4.10

What is the analog of the value iteration update (4.10) for action values, $q_{k+1}(s,a)$?

**My answer:**

$$\begin{aligned} q_{k+1}(s,a) &= \mathbb{E}\big[R_{t+1} + \gamma \max_{a'} q_k(S_{t+1},a') | S_t = s, A_t = a\big] \\ &= \sum_{s',r} p(s',r|s,a)\big[r + \gamma \max_{a'} q_k(s',a')\big] \end{aligned}$$