Exercise 3.14

Recalling Bellman Equation:

$$v_\pi(s) = \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a)[r + \gamma v_\pi(s')]$$

Now consider the center cell with

$v_\pi(s) = 0.7,$

$\gamma = 0.9,$

$\pi(a|s) = 1/4$ (The policy chooses one out of 4 directions randomly),

We will check the equation corresponding to the neighborhood cells:

$$v_\pi(s) = \frac{0.9}{4}[2.3 + 0.4 - 0.4 + 0.7]$$

$$= 0.675$$

Exercise 3.24

Recalling the value function for a policy:

$$v_*(s) = \sum_{t=0}^{\infty} \gamma^t R_t$$

The best solution after reaching A is to quickly go back A after moving to A. That takes 5 time steps. So we will have:

$$v_*(A) = \sum_{t=0}^{\infty} 10\gamma^{5t} \approx 24.419$$