Exercise 3.14

Recalling Bellman equation for state-value v_π

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

Now consider the center cell with $v_\pi(s) = 7, \quad \gamma = 0.9, \quad \pi(a|s) = \frac{1}{4} \ \ \forall a \in \mathrm{A}$

We will check if ythat this equation holds for the center state, with respect to its four neighboring states

$$\begin{aligned} v_\pi(s) &= \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r \mid s, a) \left[r + \gamma v_\pi(s')\right] \\ &= \frac{1}{4}(0 + 0.9 * 2.3) + \frac{1}{4}(0 + 0.9 * 0.4) + \frac{1}{4}[0 + 0.9 * (-0.4)] + \frac{1}{4}(0 + 0.9 * 0.7) \\ &= 0.675 \\ &\approx 0.7 \text{ (rounded to 1 decimal place)} \end{aligned}$$

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 to A. That takes 5 time steps. So we will have:

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