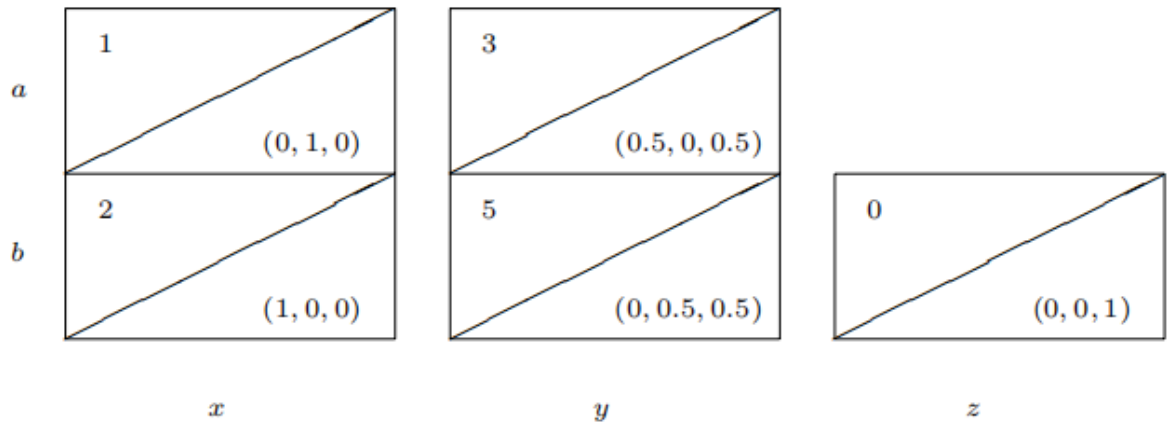
Exercise 3.7

学号	姓名
20337025	崔璨明



从$s = \{a(x) = a, a(y) = a\}, \delta = \frac{9}{10}$开始:

$s_0 = \{a(x) = a, a(y) = a\}$

Iteration 1:
The payoffs for s_0 are found from:

$$\pi(x|s_0) = 1 + \delta\pi(y|s_0)$$

$$\pi(y|s_0) = 3 + \frac{1}{2}\delta\pi(x|s_0)$$

解得:

$$\pi(x|s_0) = \frac{740}{119}, \pi(y|s_0) = \frac{690}{119}$$

改变action得到的payoffs:

$$r(x, b) + \delta\pi(x|s_0) = 2 + \frac{9}{10}\frac{740}{119} = \frac{904}{119}$$

$$r(y, b) + \delta\pi(y|s_0) = 5 + \frac{1}{2}\frac{9}{10}\frac{690}{119} = \frac{1811}{238}$$

因为$\frac{904}{119} > \frac{740}{119}$，且$\frac{1811}{238} > \frac{690}{119}$，故：

$$\hat{a}(x) = b \quad and \quad \hat{a}(y) = b$$

记为s_1

Iteration 2:
The payoffs for s_1 are found from:

$$\pi(x|s_1) = 2 + \delta\pi(x|s_1)$$

$$\pi(y|s_1) = 5 + \frac{1}{2}\delta\pi(y|s_1)$$

解得:

$$\pi(x|s_1) = 20, \pi(y|s_1) = \frac{100}{11}$$

改变action得到的payoffs:

$$r(x,a) + \delta\pi(y|s_1) = 1 + \frac{9}{10}\frac{100}{11} = \frac{101}{11}$$

$$r(y,a) + \frac{1}{2}\delta\pi(x|s_1) = 3 + \frac{9}{20} * 20 = 12$$

因为$\frac{101}{11} < 20$且$12 > \frac{100}{11}$，因此$\hat{a}(x) = b \quad and \quad \hat{a}(y) = a$，记为$s_2$

Iteration 3:
The payoffs for s_2 are found from:

$$\pi(x|s_2) = 2 + \delta\pi(x|s_2)$$

$$\pi(y|s_2) = 3 + \frac{1}{2}\delta\pi(y|s_2)$$

解得:

$$\pi(x|s_2) = 20, \pi(y|s_2) = 12$$

改变action得到的payoffs:

$$r(x,a) + \delta\pi(y|s_2) = 1 + \frac{9}{10} * 12 = \frac{118}{10}$$

$$r(y,a) + \frac{1}{2}\delta\pi(x|s_2) = 5 + \frac{9}{20} * 12 = \frac{104}{10}$$

两者都小于原来的，所以再改变策略并不会带来更好的payoffs，因此

$$s^* = s_2 = \{a(x) = b, a(y) = a\}$$