

根据:

$$V^* = \max_a Q^*(s, a)$$

$$Q^*(s, a) = \sum_{s'} P_{ss'}^a (R_{ss'}^a + \gamma V^*(s'))$$

得到相邻状态最优值关系:

$$V^*(s) = \max_a \sum_{s'} P_{ss'}^a (R_{ss'}^a + \gamma V^*(s'))$$

