**Problem 1.**

We have the Time Epochs (infinite-horizon problems). Let be the result of this game, the state space , the action space include two ways of playing , 1 means play timid, and 2 means play bold. If the state is win, we give the reward 1. If the state is lose, we give the reward -1, and if the state is draw, we give the reward 0. We should let the sum of all reward be 1, when .

The transition probability is:

And the reward function is:

The best way to win is timid-timid-timid-timid-timid-bold.

 

**Problem 2.**

We have the Time Epochs (finite-horizon problems). Let be the state of this software manufacturer, the state space , the action space include two choices , 0 means no investment in new development occurs, and 1 means the company invest in development of upgraded version of the software. When t=12, We should let the sum of all reward to be most.

The transition probability is:

And the reward function is:

If the initial state is state1, the explanation of best choice is 36.02778, and if the initial state is state 2, the explanation of best choice is 29.36111.

  

**Problem 3.**

1. We have the Time Epochs (infinite-horizon problems). Let be the state of this equipment, the state space , the action space include two choices , 0 means continue, and 1 means replace.

The transition probability is:

And the reward function is:

Our objective may be to maximize the the difference between cost, salvage and revenue per period.

2)

