

## Solution to Midterm, 6.231, 2009 Fall

### Problem 1

(a) We consider a completely connected directed graph consisting of  $n$  nodes (each represents a currency). The length of an arc from  $i$  to  $j$  is  $a_{ij} = -\ln r_{ij}$ . Now we consider the shortest path problem from some chosen node  $i$  to all other nodes  $j$ . With an arbitrage opportunity, there exists a negative cycle in the graph, and thus the optimal cost from any node  $i$  to  $j$  is not finite ( $-\infty$ ).

An alternative way is to add two artificial nodes, the origin  $s$  and the destination  $t$ , to the constructed graph. In the new graph with  $n + 2$  nodes, for any  $i, j \in \{1, 2, \dots, n\}$ , the length of the arc from  $i$  to  $j$  is  $a_{ij} = -\ln r_{ij}$ . For any  $i \in \{1, 2, \dots, n\}$ , we let  $a_{si} = a_{it} = 0$ . With an arbitrage opportunity, there exists a negative cycle in the graph, and thus the optimal cost from  $s$  to  $t$  is not finite ( $-\infty$ ).

(b) In the constructed graph with  $n$  nodes, we employ the label correcting algorithm to find a shortest path from  $i$  to another node  $j$ . To detect the existence of an arbitrage opportunity, we will find a threshold  $K$  for the value of "UPPER". Once the value of "UPPER" goes below the threshold  $K$ , we know there exists an arbitrage opportunity.

Let  $R$  be the set of all pairs of nodes,  $(i, j) \in \{1, 2, \dots, n\}^2$  such that  $i \neq j$  and  $r_{ij} > 1$ . We define the threshold  $K$  as

$$K = \sum_{(i,j) \in R} a_{ij} = - \sum_{(i,j) \in R} \ln(r_{ij}).$$

If the value "UPPER" is lower than  $K$ , it means that some arc has been visited twice by label correcting algorithm. Hence there must exist an negative cycle in the constructed graph, and we conclude that there exists an arbitrage opportunity.

We can also apply the label correcting method to the constructed graph with  $n + 2$  nodes, to find a shortest path from  $s$  to  $t$ . The threshold  $K$  also applies to this case.

### Problem 2

(a) The state for the DP problem is  $x_k$ , and the system equation is

$$x_{k+1} = h_k(p_k(x_k, u_k), w_k).$$

Thus we have,

$$J_k(x_k) = \min_{u_k \in U_k(x_k)} [g(x_k, u_k) + E_{w_k} \{J_{k+1}(h_k(p_k(x_k, u_k), w_k))\}], \quad k = 0, 1, \dots, N - 1.$$

(b) We consider a DP problem of length  $2N + 1$ . The system evolves as

$$x_0 \rightarrow y_0 \rightarrow x_1 \rightarrow y_1 \cdots, x_{N-1} \rightarrow y_{N-1} \rightarrow x_N.$$

At even stages  $2k$  ( $k = 0, 1, \dots, N$ ), the state of the DP problem is  $x_k$ ; at odd stages  $2k+1$  ( $k = 0, 1, \dots, N-1$ ), the state is  $y_k$ . At even stages  $2k$ , we can apply control  $u_k$  and the system equation is given by

$$y_k = p_k(x_k, u_k);$$

the cost function is  $g(x_k, u_k)$ . At odd stages  $2k + 1$ , there is no control available and the system evolves as

$$x_{k+1} = h_k(y_k, w_k).$$

Hence, for  $k = 0, 1, \dots, N - 1$ , the cost-to-go functions of the DP problem can be given by

$$\begin{aligned} J_k(x_k) &= \min_{u_k \in U_k(x_k)} [g(x_k, u_k) + V_k(p_k(x_k, u_k))] \\ V_k(y_k) &= E_{w_k} \{J_{k+1}(h_k(y_k, w_k))\}, \end{aligned}$$

where  $J_k(x_k)$  is the cost-to-go function at stage  $2k$ , and  $V_k(y_k)$  is the cost-to-go function at stage  $2k + 1$ .

(c) From the cost-to-go functions obtained in part b), we have

$$J_{k+1}(x_{k+1}) = \min_{u_{k+1} \in U_{k+1}(x_{k+1})} [g(x_{k+1}, u_{k+1}) + V_k(p_{k+1}(x_{k+1}, u_{k+1}))].$$

Substituting  $x_{k+1} = h_k(y_k, w_k)$  to the above equation, we have

$$J_{k+1}(h_k(y_k, w_k)) = \min_{u_{k+1} \in U_k(h_k(y_k, w_k))} [g_{k+1}(h_k(y_k, w_k), u_{k+1}) + V_{k+1}(p_{k+1}(h_k(y_k, w_k), u_{k+1}))].$$

Substituting the above equation to the cost-to-go function  $V_k(y_k) = E_{w_k} \{J_{k+1}(h_k(y_k, w_k))\}$ , finally we have

$$V_k(y_k) = E_{w_k} \left\{ \min_{u_{k+1} \in U_k(h_k(y_k, w_k))} [g_{k+1}(h_k(y_k, w_k), u_{k+1}) + V_{k+1}(p_{k+1}(h_k(y_k, w_k), u_{k+1}))] \right\}.$$

### Problem 3

(a) Let the state be the current consecutive number of days that the machine has not been maintained. Because there exists  $m$  such that  $p_m B > M$ , we can restrict the state space to  $\{0, 1, \dots, m\}$ . At each state, we have two controls, i.e., to maintain or not maintain.

At state  $i$ , if we choose to maintain, we have a cost  $M$  and the system goes to state 0; otherwise, the system moves to state  $i + 1$  with probability  $1 - p_i$ , or moves to state 0 with probability  $p_i$  and cost  $B$ .

$$J^*(i) = \min[M + \alpha J^*(0), p_i(B + \alpha J^*(0)) + \alpha(1 - p_i)J^*(i + 1)], \quad i = 0, 1, \dots, m - 1,$$

$$J^*(m) = M + \alpha J^*(0).$$

(b) Since  $p_i$  is monotonically non-decreasing in  $i$ , intuitively it is quite clear that  $J^*(i)$  is monotonically non-decreasing in  $i$ . Hence, the optimal policy is a threshold policy: maintain if and only if  $i \geq i^*$ , where  $i^*$  is the smallest integer  $i$  which satisfies,

$$p_i(B + \alpha J^*(0)) + (1 - p_i)\alpha J^*(i + 1) \geq M + \alpha J^*(0).$$

We will show that  $J^*(i)$  is monotonically nondecreasing in  $i$ . First we have

$$J^*(m - 1) = \min[M + \alpha J^*(0), p_{m-1}(B + \alpha J^*(0)) + \alpha(1 - p_{m-1})J^*(m)] \leq M + \alpha J^*(0) = J^*(m).$$

Based on the hypothesis that  $J^*(i+2) \geq J^*(i+1)$ , we will show that  $J^*(i+1) \geq J^*(i)$ .

$$\begin{aligned} J^*(i+1) &= \min[M + \alpha J^*(0), p_{i+1}(B + \alpha J^*(0)) + \alpha(1 - p_{i+1})J^*(i+2)] \\ &\geq \min[M + \alpha J^*(0), p_i(B + \alpha J^*(0)) + \alpha(1 - p_i)J^*(i+1)] \\ &= J^*(i), \end{aligned}$$

where the inequality holds because

$$\begin{aligned} &p_{i+1}(B + \alpha J^*(0)) + (1 - p_{i+1})\alpha J^*(i+2) - [p_i(B + \alpha J^*(0)) + (1 - p_i)\alpha J^*(i+1)] \\ &\geq p_{i+1}(B + \alpha J^*(0)) + (1 - p_{i+1})\alpha J^*(i+1) - [p_i(B + \alpha J^*(0)) + (1 - p_i)\alpha J^*(i+1)] \\ &= (p_{i+1} - p_i)[(B + \alpha J^*(0)) - J^*(i+1)] \\ &\geq 0, \end{aligned}$$

where the first inequality is due to the hypothesis that  $J^*(i+2) \geq J^*(i+1)$ , and the second inequality holds because  $p_{i+1} \geq p_i$  and  $B + \alpha J^*(0) \geq J^*(i+1)$ . To see why  $B + \alpha J^*(0) \geq J^*(i+1)$ , first from the cost-to-go function we have  $M + \alpha J^*(0) \geq J^*(i+1)$ . Since  $p_m B > M$ , we have  $B + \alpha J^*(0) \geq M + \alpha J^*(0) \geq J^*(i+1)$ . We have finished the proof that  $J^*(i)$  is monotonically nondecreasing in  $i$ .

(c) As for the discounted cost version, we may consider only states  $\{0, 1, \dots, m\}$ . The system starts at state 0, and it will be revisited within  $m$  stages, because we must maintain the machine at state  $m$ . Hence, state 0 is recurrent under all policies, and Assumption 7.4.1 holds. So we may apply Bellman's equation:

$$\lambda^* + h^*(i) = \min[M + h^*(0), p_i(B + h^*(0)) + (1 - p_i)h^*(i+1)], \quad i = 0, 1, \dots, m-1,$$

$$\lambda^* + h^*(m) = M + h^*(0),$$

$$h^*(0) = 0.$$

MIT OpenCourseWare  
<http://ocw.mit.edu>

6.231 Dynamic Programming and Stochastic Control  
Fall 2015

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.