Winter 2006   Homework Number 2

20/11/2006   Due 4/12/2006

For the latest see http://www.ee.technion.ac.il/~adam/GRADUATES/048913

Consider a finite (state and action) MDP.

1. Show that Markov policies suffice whenever the cost depends only on marginal distributions.

2. For a randomized Markov policy $\pi = \{\mu_0, \ldots, \mu_N\}$ define the transitsion matrix $P(\mu_k)$ and write the value of the policy using matrix notation (i.e. without using expectations).

3. Using the proof of Dynamic Programming show that deterministic policie suffice. Give conditions under which randomized policies are optimal.

4. Computer assignment. Consider the inventory problem

$$x_{k+1} = x_k + u_k - w_k \tag{0.1}$$

$$c_k(x_k, u_k, w_k) = cu_k + b[x_{k+1}]_- + h[x_{k+1}]_+ \tag{0.2}$$

where the state space is the integers, u is the control (how much we order), c_k the immediate cost at time k, c, b, h are positive, and $[x]_-$ is the negative part of x. The time horizon $N = 10$, initial stock $x_0 = 0$, and the control is limited by $0 \leq u \leq 10$. The demand w_k is i.i.d. and takes value in $[0, 15]$.

   (a) What is the number of Markov deterministic policies?

   (b) What is the number of operations (addition, multiplication, comparisons) needed to find the optimal policy by computing the values of all policies?

   (c) Choose a distribution for w and value for the other parameters. Find the optimal value and policy and display graphically. Write an analytic formula in case the policy has a simple form.