EE807 Assignment solution 1

Problem 1

(a)

- $w_k = 0$ (no disturbance)
- control constraint set $U_k(x_k) := \{u \mid 0 \le x_k + u \le 5, u : \text{integer}\}$

$$
\begin{aligned}
& 0 \le x_k + u_k \le 5 \qquad \forall k \\
\Leftrightarrow \quad & 0 \le x_{k+1} \le 5 \qquad \forall k
\end{aligned}
$$

with $x_0 = 5 \quad \Rightarrow \quad 0 \le x_k \le 5 \qquad \forall k$

$\Rightarrow \quad$ states only take the values $0, \ldots, 5$

- $N = 4$

Apply the Dynamic Programming Algorithm (DPA)

- $\underline{k = N}$

$$J_N(x_N) = 0 = J_4(x_4)$$

- $\underline{k = 3}$

$$
\begin{aligned}
J_3(x_3) &= \min_{-x_3 \le u_3 \le 5 - x_3} \left(x_3^2 + u_3^2 + J_4(x_3 + u_3)\right) \\
&= \min_{-x_3 \le u_3 \le 5 - x_3} \left(x_3^2 + u_3^2 + 0\right)
\end{aligned}
$$

$\Rightarrow \quad$ optimal control: $\underline{\underline{u_3 = \mu_3(x_3) = 0}}$

$\Rightarrow \quad \underline{\underline{J_3(x_3) = x_3^2}}$

- $\underline{k = 2}$

$$
\begin{aligned}
J_2(x_2) &= \min_{-x_2 \le u_2 \le 5 - x_2} \left(x_2^2 + u_2^2 + J_3(x_2 + u_2)\right) \\
&= \min_{-x_2 \le u_2 \le 5 - x_2} \left(2x_2^2 + 2x_2 u_2 + 2u_2^2\right)
\end{aligned}
$$

Evaluate expression for all possible x_2, u_2:

	$u_2 = -5$	-4	-3	-2	-1	0	1	2	3	4	5
$x_2 = 0$	-	-	-	-	-	0	2	8	18	32	50
1	-	-	-	-	2	2	6	14	24	42	-
2	-	-	-	8	6	8	14	24	38	-	-
3	-	-	18	14	14	18	26	38	-	-	-
4	-	32	26	24	26	32	42	-	-	-	-
5	50	42	38	38	42	50	-	-	-	-	-

Therefore, the optimal cost and policy:

x_2	$J_2(x_2)$	$\mu_2(x_2)$
0	0	0
1	2	-1 or 0
2	6	-1
3	14	-2 or -1
4	24	-2
5	38	-3 or -2

- <u>$k = 1$</u>

$$J_1(x_1) = \min_{-x_1 \le u_1 \le 5-x_1} \left(x_1^2 + u_1^2 + J_2(x_1 + u_1) \right)$$

	$u_1 = -5$	-4	-3	-2	-1	0	1	2	3	4	5
$x_1 = 0$	-	-	-	-	-	0	3	10	23	40	63
1	-	-	-	-	2	3	8	19	34	55	-
2	-	-	-	8	7	10	19	32	51	-	-
3	-	-	18	15	16	23	34	51	-	-	-
4	-	32	27	26	31	40	55	-	-	-	-
5	50	43	40	43	50	63	-	-	-	-	-

x_1	$J_1(x_1)$	$\mu_1(x_1)$
0	0	0
1	2	-1
2	7	-1
3	15	-2
4	26	-2
5	40	-3

- <u>$k = 0$</u>

$$J_0(x_0) = \min_{-x_0 \le u_0 \le 5-x_0} \left(x_0^2 + u_0^2 + J_1(x_0 + u_0) \right)$$

given: $x_0 = 5$

$$J_0(5) = \min_{-5 \le u_0 \le 0} \left(25 + u_0^2 + J_1(5 + u_0) \right)$$

x_0	-5	-4	-3	-2	-1	0
5	50	43	41	44	52	65

$$\rightarrow \quad \mu_0(x_0 = 5) = -3 \quad , \quad J_0(x_0 = 5) = 41$$

System evolution:

$x_0 = 5$	$\rightarrow$	$u_0 = -3$	$g_0 = 34$
$x_1 = 2$	$\rightarrow$	$u_1 = -1$	$g_1 = 5$
$x_2 = 1$	$\rightarrow$	$u_2 = -1$ or 0	$g_2 = 2$ or 1
$x_3 = 0$ or 1	$\rightarrow$	$u_3 = 0$	$g_3 = 0$ or 1

(b)

- $x_4 = 5$
- $x_3 + u_3 = x_4 = 5$

Apply DPA

<u>k = N</u>

$$J_4(x_4) = J_4(5) = 0$$

<u>k = 3</u>

$$u_3 = \mu_3(x_3) = 5 - x_3$$

$$J_3(x_3) = \min_{-x_3 \le u_3 \le 5 - x_3} x_3{}^2 + u_3{}^2 + J_4(5) = x_3{}^2 + (5 - x_3{}^2)$$

<u>k = 2</u>

$$J_2(x_2) = \min_{-x_2 \le u_2 \le 5 - x_2} x_2{}^2 + u_2{}^2 + J_3(x_2 + u_2)$$

	$u_2 = -5$	-4	-3	-2	-1	0	1	2	3	4	5
$x_2 = 0$	-	-	-	-	-	25	18	17	22	33	50
1	-	-	-	-	27	18	15	18	27	42	-
2	-	-	-	33	22	17	18	25	38		-
3	-	-	43	30	23	22	27	38	-	-	-
4	-	57	42	33	30	33	42	-	-	-	-
5	75	58	47	42	43	50	-	-	-	-	-

x_2	$J_{x_2}(x_2)$	$\mu_2(x_2)$
0	17	2
1	15	1
2	17	0
3	22	0
4	30	-1
5	42	-2

k = 1

	u_2 = -5	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**	**5**
$x_2 = 0$	-	-	-	-	-	17	16	21	31	46	67
1	-	-	-	-	19	16	19	27	40	59	-
2	-	-	-	25	20	21	27	38	55	-	-
3	-	-	35	28	27	31	40	55	-	-	-
4	-	49	40	37	39	46	59	-	-	-	-
5	67	56	51	51	56	67	-	-	-	-	-

x_1	$J_{x_1}(x_1)$	$\mu_1(x_1)$
0	16	1
1	16	0
2	20	-1
3	27	-1
4	37	-2
5	51	-2 or -3

k = 0

$$J_0(x_0) = \min_{-x_0 \le u_0 \le 5 - x_0} x_0{}^2 + u_0{}^2 + J_1(x_0 + u_0) = 25 + u_0{}^2 + J_1(5 + u_0)$$

	$u_0 = -5$	-4	-3	-2	-1	0
$x_0 = 5$	66	57	54	56	63	76

Optimal control: $\mu_0(x_0) = -3$

Optimal cost: $J_3(x_3) = 54$

System evolutions :

$x_0 = 5 \rightarrow u_0 = -3$

$x_1 = 2 \rightarrow u_1 = -1$

$x_2 = 1 \rightarrow u_2 = 1$

$x_3 = 2 \rightarrow u_3 = 3$

(c)

$$\sum_{k=0}^{3}(x_k^2 + u_k^2)$$

$$x_{k+1} = x_k + u_k + w_k \quad , \quad u_k \in U_k(x_k) := \{u \mid 0 \leq x_k + u \leq 5, u : \text{integer}\}$$

- As noted in a) $x_k + u_k$ is always between 0 and 5.
- In case where $x_k + u_k$ equals 0 or 5, $w_k = 0$; otherwise w_k can take values $\{-1, 1\}$. Thus x_{k+1} also takes values $0 \ldots 5$ only.

Apply DPA

- $\underline{k = N}$

$$J_4(x_4) = 0$$

- $\underline{k = 3}$

$$\begin{aligned} J_3(x_3) &= \min_{-x_3 \leq u_3 \leq 5 - x_3} \mathop{\mathrm{E}}_{w_3}\left(x_3^2 + u_3^2 + J_4(x_4)\right) \\ &= \min_{-x_3 \leq u_3 \leq 5 - x_3} \left(x_3^2 + u_3^2\right) \\ &= x_3^2 \end{aligned}$$

$$\rightarrow \quad \underline{\underline{\mu_3(x_3) = 0}} \quad , \quad \underline{\underline{J_3(x_3) = x_3^2}}$$

- $\underline{k = 2}$

$$J_2(x_2) = \min_{-x_2 \leq u_2 \leq 5 - x_2} \mathop{\mathrm{E}}_{w_2}\left(x_2^2 + u_2^2 + J_3(x_2 + u_2 + w_2)\right)$$

Case $u_2 + x_2 = 5$ or $u_2 + x_2 = 0$:

$$J_2(x_2) = \min\left(x_2^2 + u_2^2 + J_3(x_2 + u_2)\right)$$

Case $u_2 + x_2 \neq 5$ and $u_2 + x_2 \neq 0$:

$$J_2(x_2) = \min\left(x_2^2 + u_2^2 + \tfrac{1}{2}J_3(x_2 + u_2 + 1) + \tfrac{1}{2}J_3(x_2 + u_2 - 1)\right)$$

x_2	$J_2(x_2)$	$\mu_2(x_2)$
0	0	0
1	2	-1
2	7	-1
3	15	-2 or -1
4	25	-2
5	39	-2 or -3

- $\underline{k=1}$

$$J_1(x_1) = \min_{-x_1 \le u_1 \le 5-x_1} \mathop{\mathrm{E}}_{w_1} \left(x_1^2 + u_1^2 + J_2(x_1 + u_1 + w_1)\right)$$

Case $u_2 + x_2 = 5$ or $u_2 + x_2 = 0$:

$$J_1(x_1) = \min\left(x_1^2 + u_1^2 + J_2(x_1 + u_1)\right)$$

Case $u_2 + x_2 \neq 5$ and $u_2 + x_2 \neq 0$:

$$J_1(x_1) = \min\left(x_1^2 + u_1^2 + \tfrac{1}{2}J_2(x_1 + u_1 + 1) + \tfrac{1}{2}J_2(x_1 + u_1 - 1)\right)$$

x_1	$J_1(x_1)$	$\mu_1(x_1)$
0	0	0
1	2	-1
2	8	-2
3	16.5	-2
4	28.5	-3 or -2
5	42.5	-3

- $\underline{k=0}$

$$J_0(x_0) = \min_{-x_0 \le u_0 \le 5-x_0} \mathrm{E}\left(x_0^2 + u_0^2 + J_1(x_0 + u_0 + w_0)\right)$$

$$\rightarrow \quad \underline{\underline{\mu_0(x_0 = 5) = -3}} \quad , \quad \underline{\underline{J_0(x_0 = 5) = 43.25}}$$

Problem 2

The DP formulation of the problem is the following, using the notation x' for the "x" variable of the original problem, with the types of cargo as stages $k = 1, \dots, N$:

- Initial state: $x_0 = z$.
- Control variable: $u_k = x'_k \in U_k(x_k) = \mathbb{N} \cap \left[0, \frac{x_k}{w_k}\right]$. The controls are the quantities to load from type k.
- State evolution: $x_{k+1} = x_k - u_k w_k$. The state is always the remaining capacity of the vessel to be loaded with the remaining goods $k+1, \dots, N$.
- Costs: $g_k = -u_k v_k$. And the minimand is simply the negative of the value of the loaded quantity of type k, which is additive in types.

Hereby I completely captured the original problem in a dynamic programming setting. The optimal policy of loaded quantities π solves the problem

$$\begin{aligned} J_N(x_N) &= g_N(x_N) = -u_N v_N, \\ J_k(x_k) &= \min_{\pi} \{g_k(u_k, x_k) + J_{k+1}(x_{k+1})\} = \min_{\pi} \{-u_k v_k + J_{k+1}(x_k - u_k w_k)\}. \end{aligned}$$

The solution of the problem is found by the following algorithm. For states $x < \min_i \{w_i\}$, assign $J(x) = 0$. From then on in an ascending order of states, assign $J(x) = \min_{i,\ w_i \le x} \{-v_i + J(x - w_i)\}$. Thus it is sufficient to decide at each step (however close one wants to approximate continuity) which type to store into the given amount of space, taking into account how the remaining space could be optimally filled (possibly with further units of the good selected in the step). Using this method to evaluate the states, one could also account how the optimal filling just went on, but in the end can also easily rerun a "proper" dynamic programming algorithm forward for the whole capacity of the vessel to get the optimal controls in the notation above, now using the the calculated optimal values for the remaining space to decide how many units of the given type to put in (allowing zero) in descending order of $\frac{v_k}{w_k}$, thus following the formulated problem above.

Problem 3

- Let k denotes the time that the repairman just about to _x the k + 1th site.
- The total number of stages is just the number of sites, which are needed to be repaired. In particular ,we have N = n.
- The state at stage k is the closed interval [i, j], in which the sites have been repaired, and the position the repairman currently in the interval. In particular, we have S_k = [i, j], where i ≤ j, and S_0 = [s, s], S_n = [1, n]. y_k denotes the position the repairman in the interval, i.e., y_k = 0 means the repairman at i, y_k = 1 means the repairman at j. y_0 can be set to either 0 or 1.
- The control x_k in this problem is the decision we would make at time k. In particular, we have x_k = 0 means the repairman will go left, and x_k = 1 means the repairman will go right.
- The parameters ω_k of this problem are deterministic, which are the waiting cost c_i and traveling cost t_{ij} .
- The transition function S_{k+1} = [i-(1-x_k), j +x_k], and y_{k+1} = x_k. In particular, we assume that S_1 = S_0, and y_1 = y_0 means that the repairman will always firstly repair the site where he starts. (This is just an assumption.)
- The feasible set $U_k(S_k)$ is given by

$$U_k(S_k) = \begin{cases} \{1,0\}, & S_k = [i,j], \quad 1 < i \le j < n, \\ \{1\}, & S_k = [1,j], \quad 1 < j < n, \\ \{0\}, & S_k = [i,n], \quad 1 < i < n, \\ \emptyset, & S_k = [1,n] \end{cases}$$

- The cost function

$$\Lambda_k(S_k, y_k, x_k, \omega_k) = \sum_{k \notin [i-(1-x_k), j+x_k]} c_k + t_{i(1-y_k)+jy_k, (1-x_k)(i-(1-x_k))+x_k(j+x_k)}$$

In particular, we have $\Lambda_0 = \sum_{k \neq s} c_k$, and $\Lambda_0 = 0$

After identifying all the elements of this problem, we have this problem can be formulated as following dynamic programming problem

$$q_n(S_0, y_0) = \min_{\pi} \left[\sum_{k=0}^{n-1} \Lambda_k(S_k, y_k, x_k, \omega_k) + \Lambda_n(S_n, y_n, x_n, \omega_n) \right]$$

$$= \min_{x_0 \in U_0(S_0)} [\Lambda_0(S_0, y_0, x_0, \omega_0) + q_{n-1}(\Gamma(S_0, y_0, x_0, \omega_0))]$$

Problem 4

Let $\mu_A, (\mu_B)$ be the stationary control applied when in town A, (B). The control $\mu \in \{Stay, Change\}$, We can obtain the optimal stationary control by solving Bellman's equation for each of the four possible policies. Let $\mu = (\mu_A, \mu_B)$.

For $\mu^1 = (S, S)$:

$$J_A^1 = r_A + \alpha J_A^1$$
$$J_B^1 = r_B + \alpha J_B^1$$

So,

$$J_A^1 = \frac{r_A}{1-\alpha}; \qquad J_B^1 = \frac{r_B}{1-\alpha}$$

For $\mu^2 = (S, C)$:

$$J_A^2 = \frac{r_A}{1-\alpha}$$
$$J_B^2 = r_A - c + \alpha J_A^2 = -c + \frac{r_A}{1-\alpha}$$

For $\mu^3 = (C, S)$:

$$J_A^3 = -c + \frac{r_B}{1-\alpha}; \quad J_B^3 = \frac{r_B}{1-\alpha}$$

For $\mu^4 = (C, C)$:

$$J_A^4 = r_A - c + \alpha J_A^4$$
$$J_B^4 = r_B - c + \alpha J_B^4$$

Thus,

$$J_A^4 = \frac{r_A + \alpha r_B - (1+\alpha)c}{1-\alpha^2}; \qquad J_B^4 = \frac{r_B + \alpha r_A - (1+\alpha)c}{1-\alpha^2}$$

As $\alpha \to 0$, $\bar{J}^1 = \begin{bmatrix} J_A^1 \\ J_B^1 \end{bmatrix} = \begin{bmatrix} r_A \\ r_B \end{bmatrix}$ is clearly optimal. Thus, the optimal policy is for the salesmen to stay in the town he starts in. As $\alpha \to 1$, we have:

$$(1-\alpha)\bar{J}^1 = \begin{bmatrix} r_A \\ r_B \end{bmatrix}, \qquad (1-\alpha)\bar{J}^2 = \begin{bmatrix} r_A \\ r_A \end{bmatrix}$$
$$(1-\alpha)\bar{J}^3 = \begin{bmatrix} r_B \\ r_B \end{bmatrix}, \qquad (1-\alpha)\bar{J}^4 = \begin{bmatrix} r_A + r_B - c \\ r_A + r_B - c \end{bmatrix}$$

Since $c > r_A > r_B$, μ^2 is optimal. That is, the salesman should move to A and remain there.

(b) For $c=3$, $r_A = 2$, $r_B = 1$, and $\alpha = .9$:

$$J^1 = \begin{bmatrix} 20 \\ 10 \end{bmatrix}, \qquad J^2 = \begin{bmatrix} 20 \\ 17 \end{bmatrix}, \qquad J^3 = \begin{bmatrix} 7 \\ 10 \end{bmatrix}, \qquad J^4 = \begin{bmatrix} -15.26 \\ -14.74 \end{bmatrix}$$

Thus, the optimal policy is to move into A and remain there.

Problem 5

(a) State

U : a person has an umbrella

N : a person does not have an umbrella

Action

T : a person takes an umbrella to go to the other place

L : a person leaves an umbrella

$J(N) = p(W+\alpha J(U)) + \alpha(1-p)J(U)$

$= pW + \alpha J(U)$

$J(U) = p\alpha J(U) + (1-p)\min[V + \alpha J(U), \alpha J(N)]$

$= \min[(1-p)V+\alpha J(U), p\alpha J(U) + (1-p)\alpha J(N)]$

$= \min[(1-p)V+\alpha J(U), p\alpha J(U) + p\alpha(1-p)W + \alpha^2(1-p)J(U)]$

(b) for action T

$J(U) = \frac{1-p}{1-\alpha}V$

for action L

$J(U) = \frac{\alpha p(1-p)W}{1-p\alpha-\alpha^2(1-p)}$

when

$\frac{1-p}{1-\alpha}V > \frac{\alpha p(1-p)W}{1-p\alpha-\alpha^2(1-p)}$

$p < \frac{(1+\alpha)V}{(V+W)\alpha}$

the optimal policy is to leave the umbrella. Otherwise, the optimal policy is to take the umbrella.

6. (a) The states are s^i, $i=1,\dots,n$, corresponding to the worker being unemployed and being offered a salary w^i, and $\bar{s}^i$, $i=1,\dots,n$ corresponding to the worker being employed at a salary level w^i.

Bellman's equation is

$$J(s^i) = \max\left[c + \alpha\sum_{j=1}^{n} \xi_j J(s^j)\ ,\ w^i + \alpha J(\bar{s}^i)\right],\ i=1,\dots,n \qquad (1)$$
$$J(\bar{s}^i) = w^i + \alpha J(\bar{s}^i) \qquad i=1,\dots,n \qquad (2)$$

where ξ_j is the probability of an offer at salary level w^j at any one period.

From Eq. (2), we have $\quad J(\bar{s}^i) = \frac{w^i}{1-\alpha} \quad i=1,\dots,n$
so that from Eq. (1) we obtain

$$J(s^i) = \max\left[c + \alpha\sum_{j=1}^{n} \xi_j J(s^j),\ \frac{w^i}{1-\alpha}\right].$$

Thus it is optimal to accept salary w^i if

$$w^i \geq (1-\alpha)\left(c + \alpha\sum_{j=1}^{n} \xi_j J(s^j)\right).$$

The right-hand side of the above relation gives the threshold for acceptance of an offer.

(b) In this case Bellman's equation becomes

$$J(s^i) = \max\left[c + \alpha\sum_{j=1}^{n} \xi_j J(s^j),\ w^i + \alpha\left((1-p_i)J(\bar{s}^i) + p_i\sum_{j=1}^{n}\xi_j J(s^j)\right)\right] \qquad (3)$$
$$J(\bar{s}^i) = w^i + \alpha\left((1-p_i)J(\bar{s}^i) + p_i\sum_{j=1}^{n}\xi_j J(s^j)\right). \qquad (4)$$

Let us assume without loss of generality that

$$w^1 < w^2 < \cdots < w^n.$$

Let us assume further that $p_i = p$ for all i, From Eq. (4), we have

$$J(\bar{s}^i) = \frac{w^i + p\sum_{j=1}^{n}\xi_j J(s^j)}{1-\alpha(1-p)}.$$

So it follows that

$$J(\bar{\xi}^1) < J(\bar{\xi}^2) < \cdots < J(\bar{\xi}^n) \qquad (5)$$

We thus obtain that the second term in the maximization of Eq. (3) is monotonically increasing in i, implying that there is a salary threshold above which the offer is accepted.

In the case where p_i is not independent of i, salary level is not the only criterion of choice. There must be consideration for job security (the value of p_i). However, if p_i and w^i are such that Eq. (5) holds, then there still is a salary threshold above which the offer is accepted.