ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Dynamic Systems and Control
IDSC
Institut für Dynamische Systeme und Regelungstechnik

Final Exam **January 24th, 2019**

Dynamic Programming & Optimal Control (151-0563-01)

Prof. R. D'Andrea

Solutions

Exam Duration:	**150 minutes**
Number of Problems:	**4**
Permitted aids:	One A4 sheet of paper. No calculators allowed.

Problem 1 [14 points]

a) Consider the system

$$x_{k+1} = x_k + u_k + w_k, \quad k = 0, ..., N-1 \tag{1}$$

for some finite N, where $x_k, u_k, w_k \in \mathbb{R}$, and

$$w_k = 2y_{k+1} \tag{2}$$
$$y_{k+1} = y_k + \xi_k \tag{3}$$

with $y_k, \xi_k \in \mathbb{R}$, $\underset{\xi_k}{\mathrm{E}}[\xi_k] = 0$ and $\underset{\xi_k}{\mathrm{Var}}[\xi_k] \neq 0$[1].

The cost function is given by

$$\sum_{k=0}^{N-1} x_k^2 + u_k^2. \tag{4}$$

We wish to solve the optimal control problem via the Dynamic Programming Algorithm (DPA). Let the augmented state be $\tilde{x}_k := (x_k, y_k)$.

i) The augmented state's associated dynamics $\tilde{x}_{k+1} = f_k(x_k, y_k, u_k, \xi_k) =$ ________________________________

ii) $J_N(\mathrm{x}, \mathrm{y}) =$ __________

iii) $J_{N-1}(\mathrm{x}, \mathrm{y}) =$ __________

iv) Which of the following is $J_{N-2}(\mathrm{x}, \mathrm{y})$?

- ☐ $\mathrm{x}^2 + (\mathrm{x} + 2\mathrm{y})^2/2$
- ☐ $\mathrm{x}^2 + (\mathrm{x} + \mathrm{y})^2/2$
- ☐ $(\mathrm{x} + 2\mathrm{y})^2$
- ☐ None of the above.

v) Suppose $J_{k+1}(\mathrm{x}, \mathrm{y}) = \mathrm{x} + \mathrm{y}$ for some k. The optimal cost to go (as a number in $\mathbb{R}$) at time k if we are at state $x_k = 1$ is:

- ☐ = __________
- ☐ Insufficient information

vi) Suppose $J_{k+1}(\mathrm{x}, \mathrm{y}) = \mathrm{x} + \mathrm{y}$ for some k. Additionally, you know that $x_{k-1} = 1, u_{k-1} = 4$. The optimal cost to go (as a number in $\mathbb{R}$) at time k if we are at state $x_k = 1$ is:

- ☐ = __________
- ☐ Insufficient information

[1] Recall $\underset{x}{\mathrm{Var}}[x] := \underset{x}{\mathrm{E}}\left[(x - \underset{x}{\mathrm{E}}[x])^2\right]$.

b) Consider the system

$$x_{k+1} = x_k + u_k + w_k, \quad k = 0, 1, ..., N-1. \tag{5}$$

for some finite N, where $x_k, u_k, w_k \in \mathbb{R}$.

The cost function is given by

$$x_N^2 + \sum_{k=0}^{N-1} x_k^2 + u_k^2. \tag{6}$$

We are given that $x_0 = 0$, and the disturbance w_k is a random walk, that is, $\underset{w_k|w_{k-1}=\alpha}{\mathrm{E}}[w_k - w_{k-1}] = 0$ and $\underset{w_k|w_{k-1}=\alpha}{\mathrm{Var}}[w_k - w_{k-1}] = 1$ for all $\alpha \in \mathbb{R}$ and all k, with $w_{-1} = 0$.

Select True or False for each of the system dynamics formulations below such that if used by the DPA, will either yield the corresponding optimal policy or not, respectively.

i) The state is (x_k, y_k) with the following dynamics

$$x_{k+1} = x_k + u_k + y_k + \xi_k \tag{7}$$
$$y_{k+1} = y_k + \xi_k \tag{8}$$

where the disturbance $\xi_k \in \mathbb{R}$ is such that $\underset{\xi_k}{\mathrm{E}}[\xi_k] = 0$ and $\underset{\xi_k}{\mathrm{Var}}[\xi_k] = 1$, and $y_0 = 0$.

☐ True
☐ False

ii) The state is (x_k, y_k) with the following dynamics

$$x_{k+1} = x_k + u_k + y_k + 2\xi_k \tag{9}$$
$$y_{k+1} = y_k + 2\xi_k \tag{10}$$

where the disturbance $\xi_k \in \mathbb{R}$ is such that $\underset{\xi_k}{\mathrm{E}}[\xi_k] = 0$ and $\underset{\xi_k}{\mathrm{Var}}[\xi_k] = 1/4$, and $y_0 = 0$.

☐ True
☐ False

iii) The state is x_k with the following dynamics

$$x_{k+1} = x_k + u_k + y_k \tag{11}$$

where the disturbance $y_k \in \mathbb{R}$ is such that $\underset{y_k}{\mathrm{E}}[y_k] = 0$ and $\underset{y_k}{\mathrm{Var}}[y_k] = 1$.

☐ True
☐ False

iv) The state is (x_k, y_k) with the following dynamics

$$x_{k+1} = x_k + u_k + y_k \tag{12}$$
$$y_{k+1} = y_k + \xi_k \tag{13}$$

where the disturbance $\xi_k \in \mathbb{R}$ is such that $\underset{\xi_k}{\mathrm{E}}[\xi_k] = 0$ and $\underset{\xi_k}{\mathrm{Var}}[\xi_k] = 1$, and $y_0 = 0$.

☐ True
☐ False

v) The state is (x_k, y_k) with the following dynamics

$$x_{k+1} = x_k + u_k + y_k/2 + \xi_k \tag{14}$$
$$y_{k+1} = y_k + 2\xi_k \tag{15}$$

where the disturbance $\xi_k \in \mathbb{R}$ is such that $\underset{\xi_k}{\mathrm{E}}[\xi_k] = 0$ and $\underset{\xi_k}{\mathrm{Var}}[\xi_k] = 1$, and $y_0 = 0$.

☐ True
☐ False

c) Consider the system

$$x_{k+1} = x_k + u_k + w_k, \quad k = 0, 1, ..., N-1 \tag{16}$$

for some finite N, where $x_k, u_k, w_k \in \{-10, -9, ..., 10\}$, and w_k is drawn from some known probability distribution.

The cost function is given by

$$g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \tag{17}$$

for some given functions $g_k(\cdot,\cdot,\cdot),\ k = 0, ..., N-1$ and $g_N(\cdot)$. We are given that $x_0 = 0$.

i) Select the most correct statement:

☐ Solving for the optimal policy using the DPA is more computationally efficient than a brute-force approach.
☐ Solving for the optimal policy using the DPA is less computationally efficient than a brute-force approach.
☐ Solving for the optimal policy using the DPA is just as computationally efficient as a brute-force approach.

ii) If instead of knowing the probability distribution function of w_k, we only know that the disturbance w_k is zero-mean, then which of the following is most correct:

☐ The DPA can be solved in general for any functions $g_k(\cdot,\cdot,\cdot),\ k = 0, ..., N-1$ and $g_N(\cdot)$.
☐ The DPA can be solved for some functions $g_k(\cdot,\cdot,\cdot),\ k = 0, ..., N-1$ and $g_N(\cdot)$.
☐ The DPA cannot be solved for any functions $g_k(\cdot,\cdot,\cdot),\ k = 0, ..., N-1$ and $g_N(\cdot)$, as more information is required.

iii) Solving the DPA gives us the optimal cost-to-go and optimal policy for every time k.

☐ True
☐ False

Solution 1

a) i) $(x_k + u_k + 2y_k + 2\xi_k, y_k + \xi_k)$

ii) 0

iii) x^2

iv) None of the above $(\mathrm{x}^2 + 4\underset{\xi_k}{\mathrm{E}}\left[\xi_k^2\right] + (\mathrm{x}+2\mathrm{y})^2/2)$

v) Insufficient information

vi) $-17/4$

b) i) True

ii) True

iii) False

iv) False

v) True

c) i) First option

ii) Second option

iii) True

Problem 2 [15 points]

a) Consider a quantum particle whose spin can take on values of +1/2 or -1/2. Due to interactions with the environment, the transition probabilities for its spin can be modeled as follows:

$$\Pr\left(x_{k+1} = 1/2 | x_k = 1/2\right) = 0.8 \tag{18}$$
$$\Pr\left(x_{k+1} = 1/2 | x_k = -1/2\right) = 0.1 \tag{19}$$

for all $k \geq 0$ and a-priori we have

$$\Pr\left(x_0 = 1/2\right) = 0.5 \tag{20}$$

At time $k \geq 1$, a measurement $z_k \in \{A, B, C\}$ reveals some information about the state as follows:

$$\Pr\left(z_k = A | x_k = 1/2\right) = 0.8 \tag{21}$$
$$\Pr\left(z_k = B | x_k = 1/2\right) = 0.1 \tag{22}$$
$$\Pr\left(z_k = A | x_k = -1/2\right) = 0.6 \tag{23}$$
$$\Pr\left(z_k = B | x_k = -1/2\right) = 0.1 \tag{24}$$

Suppose the first two measurements are, in sequence, C, B. Construct the equivalent shortest path (SP) problem as per the Viterbi algorithm, by filling in the blanks in Figure 1 below, such that if solved, yields the maximum a-posteriori estimate of the state sequence from $k = 0$ to $k = 2$. If $\ln(\cdot)$ is required, leave your answer in the form $\pm\ln(\#.\#\#)$ (e.g. $+\ln(1.35), -\ln(1.00)$).

Hint: Recall that the maximum a-posteriori estimate is given by

$$\arg\max_{\bar{X}_0} \Pr\left(X_0 = \bar{X}_0 | Z_1 = \bar{Z}_1\right) \tag{25}$$

$$= \arg\max_{\bar{X}_0} \Pr\left(x_0 = \bar{x}_0\right) \prod_{k=1}^{N} \Pr\left(x_k = \bar{x}_k | x_{k-1} = \bar{x}_{k-1}\right) \Pr\left(z_k = \bar{z}_k | x_k = \bar{x}_k\right) \tag{26}$$

where $X_0 = (x_0, x_1, ..., x_N)$, $\bar{X}_0 = (\bar{x}_0, \bar{x}_1, ..., \bar{x}_N)$, $Z_1 = (z_1, z_2, ..., z_N)$, *and* $\bar{Z}_1 = (\bar{z}_1, \bar{z}_2, ..., \bar{z}_N)$.

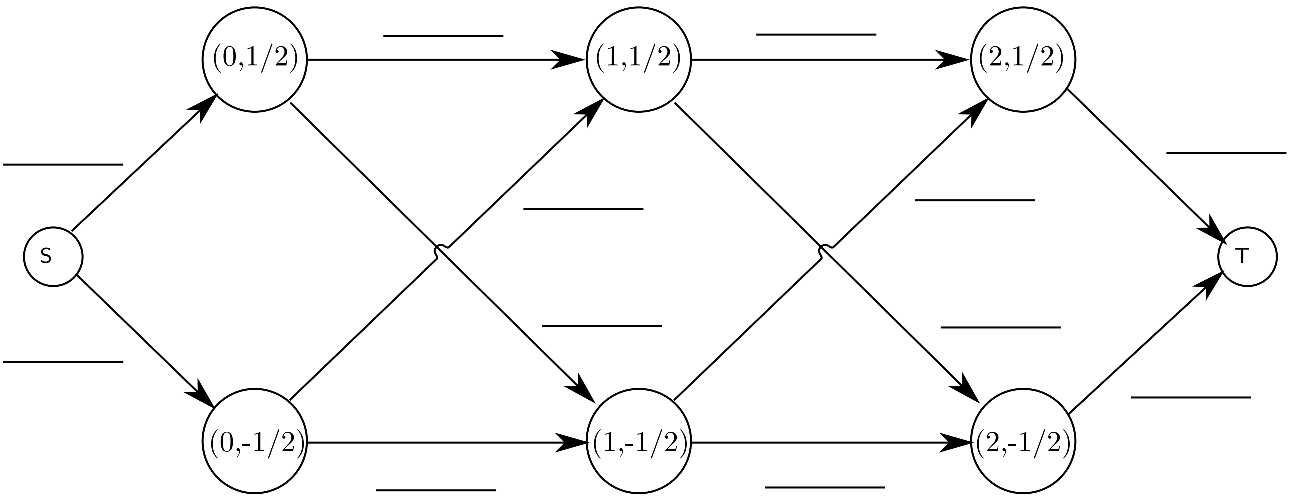

Figure 1: Estimating the quark's spin viewed as a Shortest Path problem from node S to T.

b) Consider the shortest path problem in Figure 2 below.

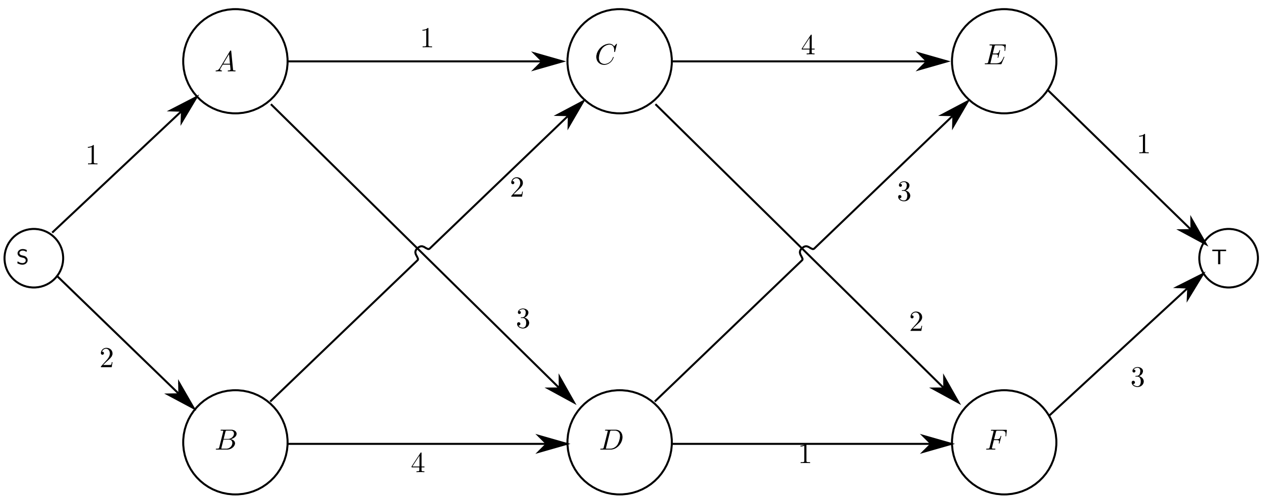

Figure 2: A Shortest Path problem from node S to T.

i) Recall that $\mathcal{V}$ is the vertex space of the graph, and $|\mathcal{V}|$ is the number of elements in the vertex space. $|\mathcal{V}| =$ __________ (leave your answer in the form # (e.g. 41, ∞)).

ii) What is the time horizon N of the equivalent Deterministic Finite State (DFS) problem? __________ (leave your answer in the form # (e.g. 41, ∞)).

iii) Consider solving the SP problem via DPA on the equivalent DFS system. Fill in the following charts (leave your answer in the form # (e.g. 41, ∞)), where $J_k(i)$ is the optimal cost to go from node i to node T in $N - k$ moves.

i	$J_{N-1}(i)$
S	________
A	________
B	________
C	________
D	________
E	________
F	________

i	$J_{N-2}(i)$
S	________
A	________
B	________
C	________
D	________
E	________
F	________

i	$J_{N-3}(i)$
S	________
A	________
B	________
C	________
D	________
E	________
F	________

iv) Consider solving the SP problem using forward DPA on the equivalent DFS system. Fill in the following charts (leave your answer in the form # (e.g. 41, ∞)), where $J_k^{\mathsf{F}}(i)$ is the optimal cost to arrive to node i from node S in k moves.

i	$J_1^{\mathsf{F}}(i)$
A	________
B	________
C	________
D	________
E	________
F	________
T	________

i	$J_2^{\mathsf{F}}(i)$
A	________
B	________
C	________
D	________
E	________
F	________
T	________

i	$J_3^{\mathsf{F}}(i)$
A	________
B	________
C	________
D	________
E	________
F	________
T	________

v) Consider solving the SP problem using the Label Correcting Algorithm (LCA) with the breadth-first (first-in/first-out) method. Fill in Table 1 below for the requested iterations.

Instructions: Recall that only one instance of a node can be in `OPEN` *at any time. If a node that is already in the* `OPEN` *bin enters the* `OPEN` *bin again, treat this node as if it would enter the* `OPEN` *bin at the current iteration. If two nodes enter the* `OPEN` *bin in the same iteration, add the one with the lowest node letter first (with* $A < B < C < ...$*).*

Example: `OPEN` *bin: B, C, D; Node exiting* `OPEN` *B (nodes entering* `OPEN`*: C, E); new* `OPEN` *bin: D, C, E; Node exiting* `OPEN` *D.*

Iteration	Remove	`OPEN`	d_S	d_A	d_B	d_C	d_D	d_E	d_F	d_T
0	—	S	0	∞	∞	∞	∞	∞	∞	∞
1										
2										
3										

Table 1: Table to be filled.

Solution 2

a) Max 4 points, every two incorrect answers or missing blanks gets minus 1 point.

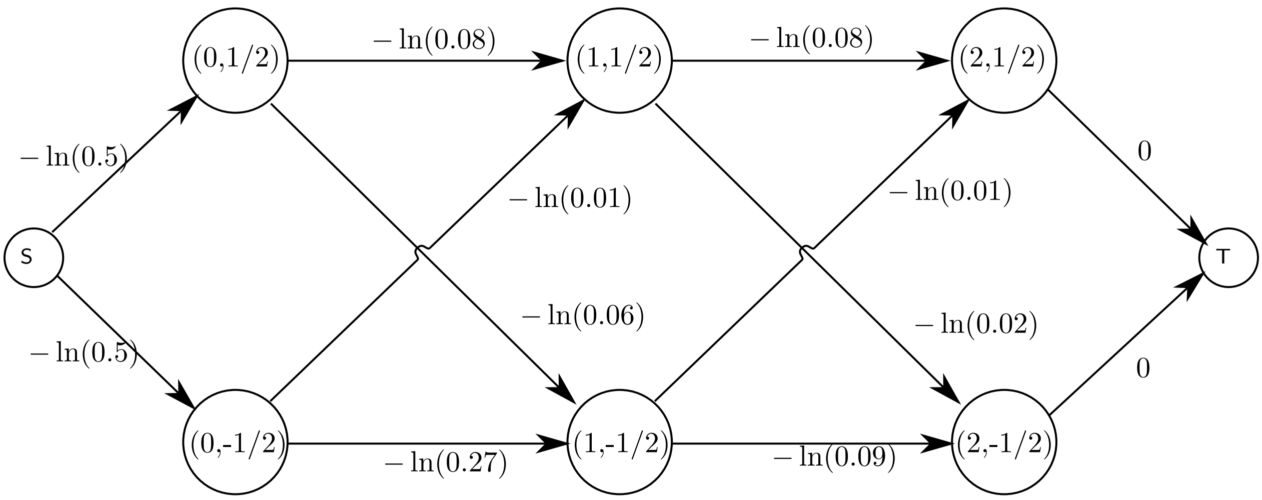

Figure 3: Estimating the quark's spin viewed as a Shortest Path problem from node S to T.

b) i) 8

ii) 7

iii) 3 points.

i	$J_{N-1}(i)$
S	∞
A	∞
B	∞
C	∞
D	∞
E	1
F	3

i	$J_{N-2}(i)$
S	∞
A	∞
B	∞
C	5
D	4
E	1
F	3

i	$J_{N-3}(i)$
S	∞
A	6
B	7
C	5
D	4
E	1
F	3

iv) 3 points.

i	$J_1^{\mathsf{F}}(i)$
A	1
B	2
C	∞
D	∞
E	∞
F	∞
T	∞

i	$J_2^{\mathsf{F}}(i)$
A	1
B	2
C	2
D	4
E	∞
F	∞
T	∞

i	$J_3^{\mathsf{F}}(i)$
A	1
B	2
C	2
D	4
E	6
F	4
T	∞

v) 3 points, 1 point per row

Iteration	Remove	`OPEN`	d_S	d_A	d_B	d_C	d_D	d_E	d_F	d_T
0	—	S	0	∞	∞	∞	∞	∞	∞	∞
1	S	A,B	0	1	2	∞	∞	∞	∞	∞
2	A	B, C, D	0	1	2	2	4	∞	∞	∞
3	B	C, D	0	1	2	2	4	∞	∞	∞

Problem 3 **[15 points]**

a) Recall the standard Stochastic Shortest Path (SSP) problem formulation: the system dynamics are

$$\begin{aligned} &x_{k+1} = w_k, \qquad x_k \in \mathcal{S}, \\ &\Pr(w_k = j | x_k = i, u_k = \mathrm{u}) = P_{ij}(\mathrm{u}), \qquad \mathrm{u} \in \mathcal{U}(i), \end{aligned} \tag{27}$$

where $\mathcal{S}$ is a finite set and $\mathcal{U}(\mathrm{x})$ is a finite set for all $\mathrm{x} \in \mathcal{S}$.

Given an initial state $i \in \mathcal{S}$, the expected closed loop cost of starting at i associated with policy $\pi = (\mu_0(\cdot), \mu_1(\cdot), \ldots, \mu_{N-1}(\cdot))$ becomes

$$J_\pi(i) = \underset{(X_1, W_0 | x_0 = i)}{\mathrm{E}} \left[\sum_{k=0}^{N-1} g(x_k, \mu_k(x_k), w_k) \right] \quad \text{subject to (27)}, \tag{28}$$

where $g(\cdot, \cdot, \cdot)$ is some given function.

The goal is to construct an optimal policy π^* such that for all $i \in \mathcal{S}$,

$$\pi^* = \arg\min_{\pi \in \Pi} J_\pi(i),$$

and explore what happens as N goes to infinity. Furthermore, two assumptions on the problem data are made:

Assumption .1 *There exists a cost-free termination state, which we designate as state* 0*. In particular, there are* $n+1$ *states with* $\mathcal{S} = \{0, 1, \ldots, n\}$*, where*

$$P_{00}(u) = 1 \text{ and } g(0, \mathrm{u}, 0) = 0, \quad \forall \mathrm{u} \in \mathcal{U}(0).$$

Assumption .2 *There exists at least one proper policy* $\mu \in \Pi$*. Furthermore, for every improper policy* μ'*, the corresponding cost function* $J_{\mu'}(i)$ *is infinity for at least one state* $i \in \mathcal{S}$*.*

Answer the following True or False questions

i) We can use value iteration to approximately solve the associated Bellman's Equation of the above problem.

ii) If the above problem has a 100 dimensional state space and a 5 dimensional control space, and we use policy iteration to solve it, the most time consuming part is solving systems of linear equations.

iii) The optimal policy π^* is always time-invariant.

iv) It is possible that the above problem has more than one termination state and still has a unique optimal cost.

v) In the above problem, it is possible that applying asynchronous policy iteration leads to the optimal solution, even if at each recursion only policies for a part of the states are updated.

vi) The term $x_k^2 + (1/k)^2$, where k is the stage number, is a possible choice for the stage cost for the above problem.

vii) Are the following matrices valid probability transition matrices (including the terminal state)? $P \in \mathbb{R}^{3\times 3}$ is the probability transition matrix under a policy μ, whose $(i,j)^{\text{th}}$ entry is $P_{ij}(\mu(i))$ where $i, j \in \mathcal{S}$.
Matrix 1:

$$P = \begin{bmatrix} 0.3 & 0.8 & 0 \\ 0.7 & 0.2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{29}$$

Matrix 2:

$$P = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.6 & 0.4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{30}$$

viii) When applying policy iteration for the above problem, it is possible that a non-optimal policy is visited twice.

b) Consider the following discounted problem represented in Figure 4, where at any state $i \in \{1, 2\}$, the control action u can either be A or B.

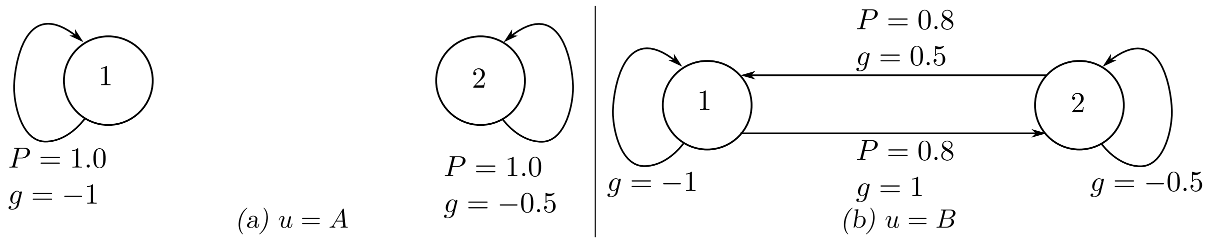

Figure 4: Probability transition graph, with the associated probabilities and costs denoted on each arc, that is,

$$P_{11}(A) = 1$$
$$P_{21}(B) = 0.8$$
$$P_{12}(B) = 0.8$$

$$g(1, u, 1) = -1, u = A, B$$
$$g(1, B, 2) = 1$$
$$g(2, B, 1) = 0.5$$
$$g(2, u, 2) = -0.5, u = A, B$$

The cost function is

$$\lim_{N\to\infty} \sum_{k=0}^{N-1} \alpha^k g(x_k, u_k, w_k) \tag{31}$$

where $\alpha \in (0, 1)$ is the discount factor.

i) $P_{22}(B) =$ (leave your answer in the form #.# e.g. 0.2).

ii) Let $q(i, \mathrm{u}) := \underset{(w|x=i,u=\mathrm{u})}{\mathrm{E}} [g(x, u, w)]$, then $q(1, B) =$ (leave your answer in the form #.# e.g. 0.2).

For questions iii) to vi), assume $\alpha = 0.5$.

iii) The associated Bellman Equation is

☐

$$J(1) = \min\{-1 + J(1),\ 0.6 + 0.2J(1) + 0.8J(2)\}$$
$$J(2) = \min\{-0.5 + J(2),\ 0.3 + 0.2J(2) + 0.8J(1)\}$$

☐

$$J(1) = \min\{-1 + 0.5J(1),\ 0.6 + 0.2J(1) + 0.8J(2)\}$$
$$J(2) = \min\{-0.5 + 0.5J(2),\ 0.3 + 0.2J(2) + 0.8J(1)\}$$

☐

$$J(1) = \min\{-1 + 0.5J(1),\ 0.6 + 0.1J(1) + 0.4J(2)\}$$
$$J(2) = \min\{-0.5 + 0.5J(2),\ 0.3 + 0.1J(2) + 0.4J(1)\}$$

☐

$$J(1) = \min\{-1 + 0.5J(1),\ 1 + 0.1J(1) + 0.4J(2)\}$$
$$J(2) = \min\{-0.5 + 0.5J(2),\ 0.5 + 0.1J(2) + 0.4J(1)\}$$

☐ None of the above

iv) Let $\mu(1) = A$, $\mu(2) = A$. The associated expected infinite horizon closed loop cost $J_\mu(1) =$ and $J_\mu(2) =$ (leave your answer in the form #.# e.g. 0.2).

v) Construct the probability transition graph of the auxiliary SSP for the control action A , by filling in the blanks in Figure 5 (leave your answer in the form #.# e.g. 0.2).

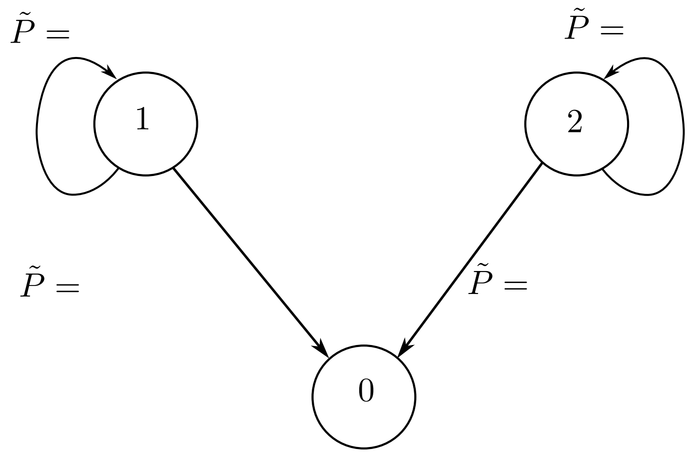

Figure 5: Probability transition graph of the control action A for the auxiliary SSP of the discounted problem represented in Figure 4

vi) An optimal policy $\mu(1) =$ and $\mu(2) =$ (fill in with A, B or both)

vii) When $\alpha = 0.8$, an optimal policy is $\mu(1) = A$, $\mu(2) = B$. Then the optimal policy for $\alpha = 0.95$ is

☐ $\mu(1) = A$, $\mu(2) = A$.
☐ $\mu(1) = A$, $\mu(2) = B$.
☐ $\mu(1) = B$, $\mu(2) = A$.
☐ $\mu(1) = B$, $\mu(2) = B$.

Solution 3

a) i) True

ii) True

iii) True

iv) True

v) True

vi) False

vii) False, True

viii) False

One point each

b) i) 0.2.

1

ii) 0.6

1

iii) The third option

1

iv) -2, -1

1

v) $\tilde{P}_{11}(A) = 0.5$, $\tilde{P}_{10}(A) = 0.5$, $\tilde{P}_{22}(A) = 0.5$, $\tilde{P}_{20}(A) = 0.5$

1 point, if there is one incorrect or missing blank, 0 points

vi) A, A

1

vii) $\mu(1) = A$, $\mu(2) = B$.

1

Problem 4 [21 points]

a) Consider the standard deterministic continuous time optimal control problem, that is, a system with dynamics

$$\dot{x}(t) = f(x(t), u(t)), \qquad 0 \leq t \leq T \tag{32}$$

and cost function

$$h(x(T)) + \int_0^T g(x(\tau), u(\tau))\mathrm{d}\tau \tag{33}$$

Answer the following true or false questions.

i) If there exists an optimal control law, the associated cost-to-go function always satisfies the Hamilton Jacobi Bellman equation.

Assume the cost function has the following form for the next six questions (question (ii) to (vii))

$$\int_0^T x(\tau)^2\mathrm{d}\tau, \tag{34}$$

where $x(t) \in \mathbb{R}$ and $u(t) \in \mathbb{R}$ for all $t \in [0, T]$.

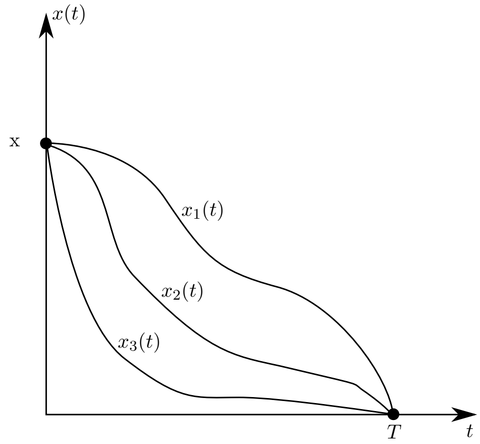

Figure 6: Three resulting state trajectories from three feedback control laws. Trajectory $x_i(t)$, $0 \leq t \leq T$, is the resulting trajectory when the feedback law $\mu_i(\cdot, \cdot)$ is applied to the system, $i = 1, 2, 3$.

Consider three feedback laws $\mu_1(\cdot, \cdot)$, $\mu_2(\cdot, \cdot)$ and $\mu_3(\cdot, \cdot)$, which result in trajectories $x_1(t), x_2(t)$ and $x_3(t)$, respectively, for the initial condition $x(0) = \mathrm{x}$. The state trajectories are also shown in Fig. 6.

Answer whether the following three statements are true or false:

ii) $\mu_3(\cdot, \cdot)$ is definitely an optimal control law.

iii) $\mu_2(\cdot, \cdot)$ yields a smaller cost than $\mu_1(\cdot, \cdot)$ for $x(0)$ starting at x.

iv) It is NOT possible that $\mu_1(\cdot, \cdot)$ is an optimal control law.

Assume there exists an optimal trajectory for a specific initial condition $x(0) = \mathrm{x}$. We solved the corresponding optimal control problem via Pontryagin's Minimum Principle and obtained three input trajectories $u_1(t)$, $u_2(t)$, $u_3(t)$, $0 \leq t \leq T$ and their corresponding state trajectories $x_1(t)$, $x_2(t)$, $x_3(t)$. Their corresponding state trajectories are shown in Fig. 7.

Answer whether the following three statements are true or false:

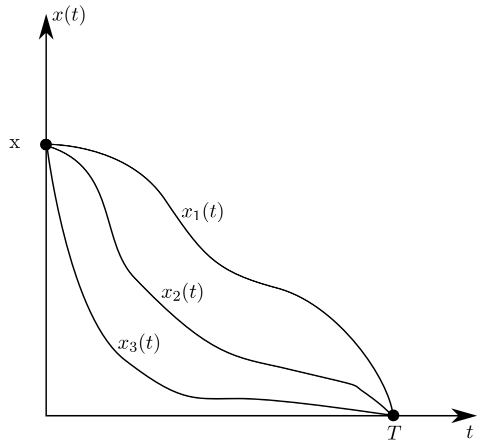

Figure 7: Three resulting state trajectories solved using Pontryagin's Minimum Principle. Trajectory $x_i(t)$, $0 \leq t \leq T$, is the resulting trajectory when $u_i(t)$ is applied to the system, $i = 1, 2, 3$.

v) $u_2(t)$, $0 \leq t \leq T$ is a feedback control law.

vi) $u_1(t)$, $0 \leq t \leq T$ is an optimal control trajectory.

vii) $u_3(t)$, $0 \leq t \leq T$ is an optimal control trajectory.

b) Consider a continuous-time system

$$\dot{x}(t) = u(t), \qquad 0 \leq t \leq T \tag{35}$$

where $x(t) \in \mathbb{R}^3$ and $u(t) \in \mathbb{R}^3$ for all $t \in [0, T]$. Furthermore,

$$x_1(0) = \text{free}, \quad x_1(T) = \text{free}, \tag{36}$$
$$x_2(0) = 0, \quad x_2(T) = 1, \tag{37}$$
$$x_3(0) = 0, \quad x_3(T) = \text{free}. \tag{38}$$

where $x_j(t)$ denotes the j-th entry of the vector $x(t)$. The cost function is

$$x_1(0)^2 + x_1(T)^2 + x_3(T)^2 + \int_0^T x_1(\tau)^2 + x_2(\tau)^2 + x_3(\tau)^2 \mathrm{d}\tau \tag{39}$$

We wish to solve this problem via Pontryagin's Minimum Principle. Answer the following questions:

i) State the number of ODEs to be solved (as an example, $\dot{x}(t) = f(x(t))$, $0 \leq t \leq T$, $x(0) = \mathrm{x}$, where $x(t) \in \mathbb{R}^2$, is considered as 2 ODEs): ____________

ii) State the number of boundary conditions for the ODEs: ____________

iii) The Hamiltonian $H(\mathrm{x}, \mathrm{u}, \mathrm{p})$, where $\mathrm{x} \in \mathbb{R}^3$, $\mathrm{u} \in \mathbb{R}^3$, $\mathrm{p} \in \mathbb{R}^3$, for this problem is

- ☐ $\mathrm{x}^\top \mathrm{x} + \mathrm{p}^\top \mathrm{u}$
- ☐ $\mathrm{x}^\top \mathrm{x} + \mathrm{p}^\top \mathrm{x}$
- ☐ $x_1(\tau)^2 + x_2(\tau)^2 + x_3(\tau)^2 + p(\tau)u(\tau)$
- ☐ None of the above

Answer whether the following equations are boundary conditions for the ODEs of this optimal control problem:

iv) $x_2(0) = 0$

v) $x_1(T) = \text{free}$

vi) $p_1(0) = 2x_1(0)$

vii) $p_1(T) = 2x_1(T)$

viii) $p_3(T) = 0$

c) In optics, the principle of least time is the principle that a ray of light always takes the path that can be traversed in the least time between two points in space. We explore a simplified 2D problem: For a particular medium (see Fig. 8), a ray of light is constrained to travel on the x-y plane, and its speed is given by $c(x, y) = x$. The objective is to find the path that the light traverses from point A $(1, 1)$ to point B $(2, 2)$ in this medium. Assume

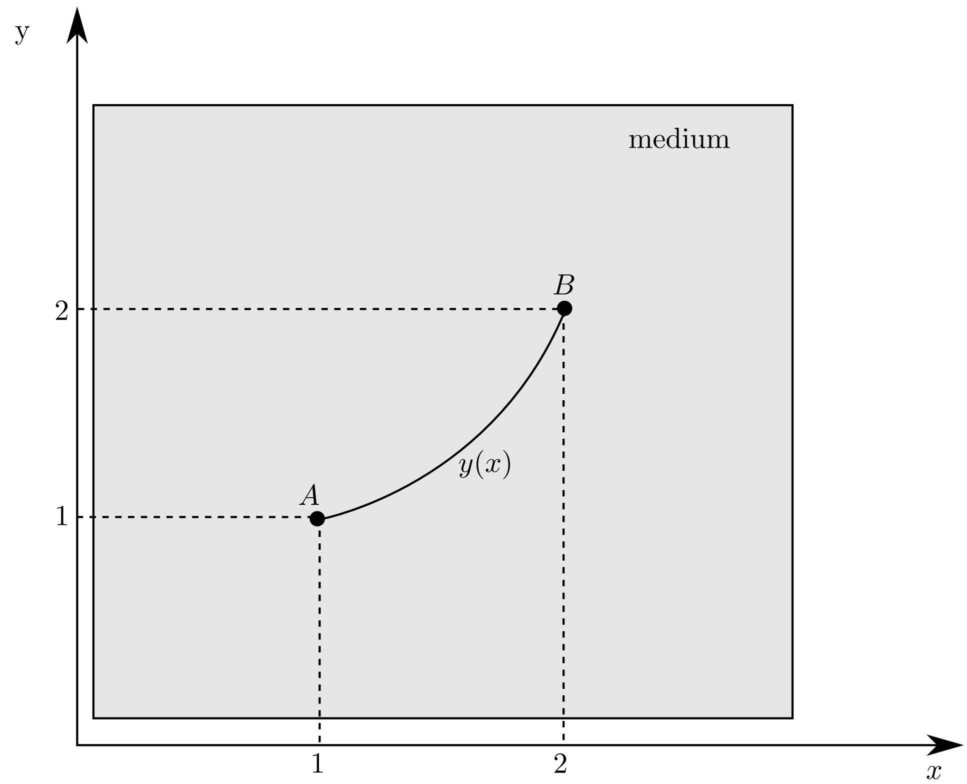

Figure 8: An example of the path $(x, y(x))$ that the light traverses in a medium

the path can be parametrized by $(x, y(x))$, where $y(x)$ is continuously differentiable with respect to x. The time that the light takes to travel from A to B can be written as

$$\int_1^2 \frac{\sqrt{\dot{y}(x)^2 + 1}}{x} \mathrm{d}x, \tag{40}$$

where $\dot{y}(x) := \mathrm{d}y(x)/\mathrm{d}x$. We can formulate this problem as a deterministic continuous time optimal control problem and solve it via Pontryagin's Minimum Principle. We treat x as the time, $y(x)$ as the state, and $\dot{y}(x)$ as the control input $u(x)$ in our formulation.

i) State the dynamics of this system $\dot{y}(x) = f(y(x), u(x))$, $x_0 \le x \le x_1$ with problem data substituted in. Specifically, replace the function $f(y(x), u(x))$ with the correct expression and x_0, x_1 with the correct numbers for this problem.

$f(y(x), u(x)) =$ ______________________ $x_0 =$ ___________ $x_1 =$ ___________

ii) The cost function in the standard formulation is

$$h(y(2)) + \int_1^2 g(y(\tau), u(\tau), \tau)\mathrm{d}\tau$$

Replace $h(y(2))$ and $g(y(\tau), u(\tau), \tau)$ with the correct expressions for this problem.
$h(y(2)) =$ ____________________
$g(y(\tau), u(\tau), \tau) =$ ____________________

iii) State the ODEs and their boundary conditions to be solved if we were to apply Pontryagin's Minimum Principle.

____

____

iv) State the Hamiltonian $H(\mathrm{y}, \mathrm{u}, \mathrm{p}, x)$ for this problem

$H(\mathrm{y}, \mathrm{u}, \mathrm{p}, x) =$ ____

v) State the minimizing $u(x)$ as a function of $y(x), p(x), x$, where

$$u(x) = \arg\min_{\mathrm{u} \in \mathcal{U}} H(y(x), \mathrm{u}, p(x), x)$$

$u(x) =$ ____________________

Solution 4

a)
- i) False
- ii) False
- iii) True
- iv) True
- v) False
- vi) False
- vii) True

1 point each

b)
- i) 6
- ii) 6
- iii) The first option
- iv) True
- v) False
- vi) False
- vii) True
- viii) False

1 point each

c)
- i) $\dot{y}(x) = u(x)$, $1 \leq x \leq 2$.

1 point

- ii) $h(y(2)) = 0$, $g(y(\tau), u(\tau), \tau)) = \frac{\sqrt{u(\tau)^2+1}}{\tau}$

2 points

- iii)

$$\dot{y}(x) = u(x), \quad 1 \leq x \leq 2, \quad y(1) = 1, \quad y(2) = 2. \tag{41}$$
$$\dot{p}(x) = 0 \tag{42}$$

Alternatively, by augmenting the state,

$$\dot{y}(x) = u(x), \quad 1 \leq x \leq 2, \quad y(1) = 1, \quad y(2) = 2, \tag{43}$$
$$\dot{z}(x) = 1, \quad z(1) = 1, \quad z(2) = 2, \tag{44}$$
$$\dot{p}_1(x) = 0, \tag{45}$$
$$\dot{p}_2(x) = \frac{\sqrt{u(x)^2+1}}{x^2} \tag{46}$$

1 point

iv) The Hamiltonian of this problem

$$H(\mathrm{y}, \mathrm{u}, \mathrm{p}, x) = \frac{\sqrt{1 + \mathrm{u}^2}}{x} + \mathrm{pu}$$

Alternatively, by augmenting the state,

$$H(\mathrm{y}, \mathrm{u}, \mathrm{p}, x) = \frac{\sqrt{1 + \mathrm{u}^2}}{x} + \mathrm{p}_1 \mathrm{u} + \mathrm{p}_2,$$

where $\mathrm{p} = (\mathrm{p}_1, \mathrm{p}_2)$.

1 point

v) To find the minimizing $u(x)$:

$$\begin{aligned} \frac{\partial H(y(x), \mathrm{u}, p(x), x)}{\partial \mathrm{u}} &:= 0 \\ \frac{u(x)}{x\sqrt{1 + u(x)^2}} + p(x) &= 0 \\ \frac{u(x)}{\sqrt{1 + u(x)^2}} &= -xp(x). \end{aligned} \tag{47}$$

which yields

$$u(x) = \pm \sqrt{\frac{x^2 p(x)^2}{1 - x^2 p(x)^2}}.$$

1 point till here

Sufficient condition:

$$\frac{\partial^2 H(y(x), \mathrm{u}, p(x), x)}{\partial \mathrm{u}^2} = \frac{1}{x(1 + u^2)^{\frac{3}{2}}} > 0, \quad \forall u(x)$$

Furthermore,

$$\begin{aligned} \dot{p}(x) &= -\left. \frac{\partial H(\mathrm{y}, \mathrm{u}, \mathrm{p}, x)}{\partial \mathrm{y}} \right|_{y(x), u(x), p(x)} \\ &= 0. \end{aligned}$$

$p(x)$ is thus a constant. Since x is always positive, combining (47) it can be seen that $u(x)$ must have the same sign over the whole interval, therefore

$$u(x) = -\frac{xp(x)}{\sqrt{1 - x^2 p(x)^2}}. \tag{48}$$

1 point bonus for correct $u(x)$