February 2022 Admissions

AY2022 Examination Questions for the Graduate School of Information Science and Engineering, Ritsumeikan University (Master's Program)

Major in Information Science and Engineering
Information Science and Engineering Course

【How to answer questions】

Please follow the instructions below and answer the questions.

<u>Choose two questions from the common subjects and choose either the Computer Science section or the Human Information Science section.</u>

In case choosing the Computer Science section, answer three questions from question ④~⑨.

In case choosing the Human Information Science section, choose one question either ⑩ or ⑪.

There will be two blank answer sheets in case choosing the Human Information Science section.

<table>
<tr><td>Common Subjects</td><td colspan="2">①Linear Algebra
②Probability and Statistics
③Data Structure and Algorithms</td></tr>
<tr><td rowspan="2">Specialized Subjects</td><td>Computer Science</td><td>④Computer Architecture
⑤Operating System
⑥Software Engineering
⑦Computer Networks
⑧Databases
⑨Artificial Intelligence</td></tr>
<tr><td>Human Information Science</td><td>⑩Image Processing
⑪Artificial Intelligence</td></tr>
</table>

【Examination time】

9:30-11:30 (120minutes)

※Leaving the examination venue is not allowed during the examination time.

※In case you feel sick or need to go to the bathroom, let examination supervisors know by raising your hand.

【Notes】

(1) Use one answer sheet for one question.

(2) Fill out your examination number and name for all the answer sheets. Also, make sure to fill out all the other necessary sections such as the questions number column.

(3) Do not remove the staple of your answer sheets.

(4) Answer sheets with no names will be invalid. Do not take the question sheets and answer sheets with you after the examination.

Common Subjects

①Linear Algebra
②Probability and Statistics
③Data Structure and Algorithms

Choose two questions from the above.

Master's Program, Information Science and Engineering Course,
Graduate School of Information Science and Engineering,
Ritsumeikan University

Common Subjects① Linear Algebra

Describe the process, not just the answer, for the following questions.

Question 1. Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 3 & 4 & 5 & 6 \\ 0 & 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix}$. Answer the following questions.

(1) Find the inverse matrix A^{-1} of A.

(2) Find the maximum and minimum eigenvalues of A^{-1}, which is the inverse matrix of A.

(3) Find the eigenvector corresponding to the maximum eigenvalue of A^{-1}, which is the inverse matrix of A.

Question 2. Find the ranks of the following matrices. (Please also show the reasoning behind your decision)

(1) $\begin{bmatrix} 2 & 1 & 1 & 8 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

(2) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

(3) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

(4) $\begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

This page is blank

Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

---

Common Subjects② Probability and Statistics

Answer all the questions below with the derivation process.

Question 1. Find each constant C_i $(i = 1,2, \dots 8)$, the expectation and the variance of the following probability functions.

(1) Probability distribution : $P(X = k) = \begin{cases} C_1, & (k = 0,\ \pm 1,\ \pm 2), \\ 0, & \text{(Otherwise)} \end{cases}$

(2) Probability distribution : $P(X = k) = \begin{cases} C_2, & (k = 9998,\ 9999,\ 10000,\ 10001,\ 10002), \\ 0, & \text{(Otherwise)} \end{cases}$

(3) Probability distribution : $P(X = k) = \begin{cases} C_3, & (k = 0,\ \pm 100,\ \pm 200), \\ 0, & \text{(Otherwise)} \end{cases}$

(4) Probability distribution : $P(X = k) = \begin{cases} C_4 \cdot k^2, & (k = 0,\ \pm 1,\ \pm 2), \\ 0, & \text{(Otherwise)} \end{cases}$

(5) Probability density distribution : $f(x) = \begin{cases} C_5, & (|x| \le 2), \\ 0, & \text{(Otherwise)} \end{cases}$

(6) Probability density distribution : $f(x) = \begin{cases} C_6 \cdot x^2, & (|x| \le 2), \\ 0, & \text{(Otherwise)} \end{cases}$

(7) Probability density distribution : $f(x) = C_7 \cdot \exp\left(-\frac{1}{4}x^2 + x\right),\ (x \in \mathrm{R})$

(8) Probability density distribution : $f(x, y) = C_8 \cdot \exp\left(-\frac{1}{4}x^2 - y^2 + x - 2y - 2\right),\ (x, y \in \mathrm{R})$

For the two-variable function, in addition to the constant C_8 ,give the covariance, the expectation and the variance of each variable.

Question 2. Five samples are obtained for two-dimensional data (X, Y) as follows:

$$(X, Y) = (1, 1), (2, 3), (3, 2), (4, 5), (5,4) .$$

Find the sample average and variance of each X and Y,and the covariance and correlation coefficient between X and Y.

This page is blank

Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

Common Subjects③ Data Structure and Algorithms

Answer all the questions below.

Question 1.

Answer the following questions about binary search trees. It is assumed that the binary search tree satisfies the following requirements.

The value of the left child of point v ＜ The value of point v ＜ The value of the right child of point v

(1) The number sequence {21, 5, 37, 28, 8, 19, 3} is inserted into the empty binary search tree in this order. Show the diagram of this binary search tree.

(2) Show the diagram of the binary search tree after deleting element {21} from the binary search tree in (1).

(3) A binary search tree of height 2 was created by re-ordering the elements of the sequence of numbers in (1). Show a diagram of this binary search tree.

Question 2.

Answer the following questions about heap. In addition, each element of the heap must satisfy the following requirements.

The value of point v is always less than or equal to the value assigned to the children of point v.

(1) A sequence of numbers {41, 25, 44, 50, 22, 31} was inserted into an empty heap in this order. Show a diagram of this heap.

(2) Show the diagram of the heap after the smallest element is deleted from the heap in (1).

Question 3.

For the time complexity of an algorithm on data of size n, give the words that fit in (a)-(d) in the following sentences.

The time complexity of the linear search algorithm is O((a)). The time complexity of the binary search, which divides the data to be searched into halves, is O((b)). For the heap, the time complexity required to reconstruct it by adding one element is O((c)). Also, the amount of time required to delete the smallest element and reassemble the heap is O((d)).

Specialized Subjects

Computer Science

④Computer Architecture
⑤Operating System
⑥Software Engineering
⑦Computer Networks
⑧Databases
⑨Artificial Intelligence

In case choosing the Computer Science section, answer three questions from question ④～⑨

Master’s Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

Computer Science④　Computer Architecture

Answer all the questions below.

Question 1.

Explain the difference between a direct-mapped cache and a fully-associative cache in terms of how to map a memory block to a cache block. Also explain the difference between the two caches in terms of cache hit ratios and hardware complexity.

Question 2.

Explain what control hazard is by using an example. Also explain how to eliminate or reduce the effect of control hazard.

Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

Computer Science⑤ Operating System

Answer all the questions below.

Question:
Summarize and explain the related technologies of virtual memory systems commonly used now by using all the technical terms below correctly. Make a complete explanation <u>using only one side of the answer sheet.</u>

[Technical terms]
TLB, FIFO, LRU, OPT, replacement, virtual address, virtual memory, memory management, locality, reference string, real address, main memory, segmentation, page table, secondary memory, contiguous allocation, fetch, paging, auxiliary memory, non-contiguous allocation

This page is blank

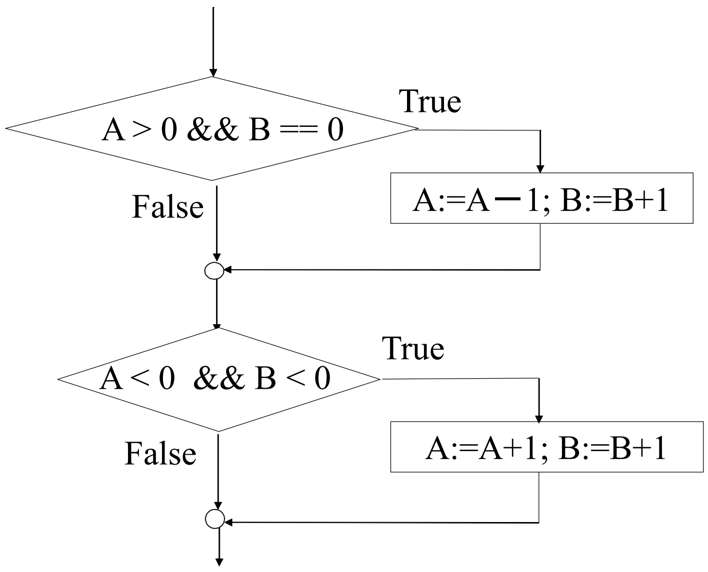
Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

Computer Science⑥ Software Engineering

Answer all the questions below.

Question 1:

In the following flowchart, "&&" is the logical AND operation. A and B are integer type variables and their values should be either −1, 0 or 1. Show the minimum test cases that completely satisfy branch coverage criteria.



Question 2: Describe two advantages of throwaway prototyping comparing with evolutional prototyping.

Question 3: Explain the disadvantages of stamp coupling in software modular design.

Question 4: Explain the meaning of aggregation in an object-oriented approach.

Question 5: Explain two purposes of software testing.

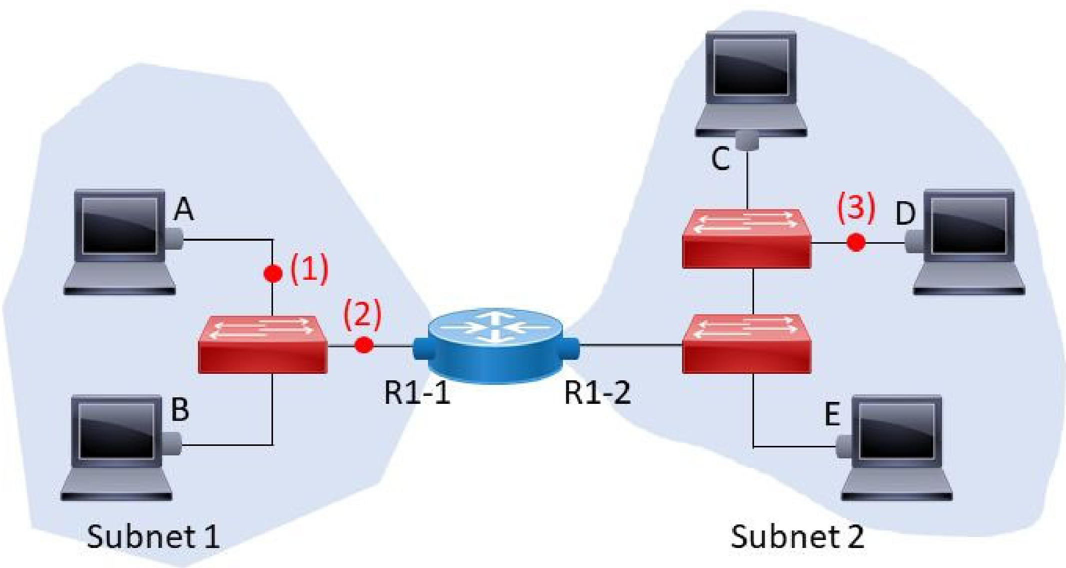
Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

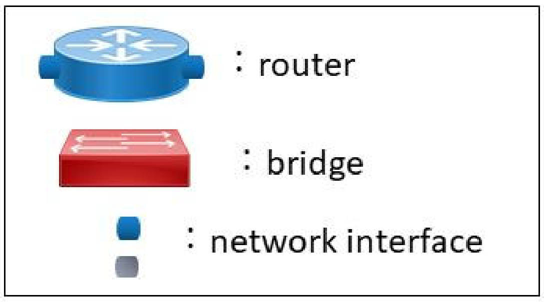
Computer Science⑦ Computer Networks ＊This question consists of 2pages.

Answer all the questions below.

Question 1.

Consider an IP datagram being sent from A to D using Ethernet as the data link-layer protocol in all links in the figure below. What are the source MAC address, destination MAC address, source IP address, and destination IP address of the IP datagram encapsulated within the Ethernet frame at points (1), (2), and (3) in the below example for a datagram going from A to D. Write all four addresses at each of the three points.





Node (network interface)	IP address	MAC address
R1-1	192.168.1.1	20:FF:3A:BC:01:4E
A	192.168.1.2	AA:12:F3:5C:01:BC
B	192.168.1.3	BB:D4:E1:A2:97:F0
R1-2	192.168.2.1	20:FF:3A:BD:12:3A
C	192.168.2.2	CC:89:34:E7:01:3B
D	192.168.2.3	DD:26:14:B7:01:3F
E	192.168.2.4	EE:B2:43:C4:BB:BC

Question 2.

You rip a 3 minute long stereo music track by sampling it at 8kHz with 16bits per channel and store it in a file in raw format. How long would it take to send the file over a 1Mbps communication link? Select the correct answer.

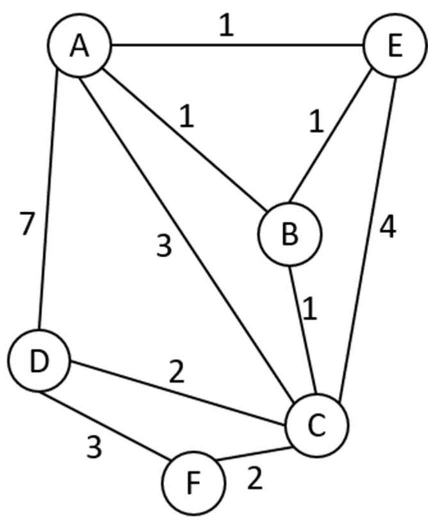
(1) About 23 seconds.
(2) About 46 seconds.
(3) About 3 minutes.
(4) About 6 minutes.
(5) It depends on the volume setting.

Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

＊This question consists of 2pages.

Question 3.

Consider the network shown below and assume that each node initially knows the costs to each of its neighbors. Use the distance vector algorithm and complete the following distance table for node C after the algorithm has converged. Write the costs for (1) through (14). Make sure to consider the Poison Reverse algorithm in your calculation. In other words, some of the entries might have a cost of ∞.



Distance Vector table at node C		Cost to destination node					
		A	B	C	D	E	F
Distance Vector from neighbor	A	(1)	(2)	(3)	(4)	(5)	(6)
	B	1	0	∞	∞	1	∞
	D	∞	∞	∞	0	∞	3
	E	(7)	(8)	(9)	(10)	(11)	(12)
	F	∞	∞	∞	(13)	∞	(14)

Master's Program, Information Science and Engineering Course,
Graduate School of Information Science and Engineering,
Ritsumeikan University

Computer Science⑧ Databases

Answer all the questions below.

Question 1. Choose a correct word from the choices and answer with a symbol (a) to (s) to fill in each of the numbered brackets, for the explanation on the relational database design below. (Brackets with the same number are filled with the same word, while brackets with different numbers may also be filled with the same word.)

Consider a database to manage the vaccination history of citizens. In case of adding a "Date_of_Vaccination" to the entity-relationship diagram in Figure 1, this data should be an attribute of [①]. When the relationship "Injected" is transformed to a relation, the primary key of this relation will be [②]. For a relation transformed from the entity "Vaccine", there are functional dependencies that [③] uniquely determines [④] and [④] uniquely determines [⑤]. Therefore, this relation satisfies the condition up to [⑥] normal form. (Note that a citizen is vaccinated multiple times with different lot numbers of vaccine, and the same lot number of vaccine is injected to multiple citizens.)

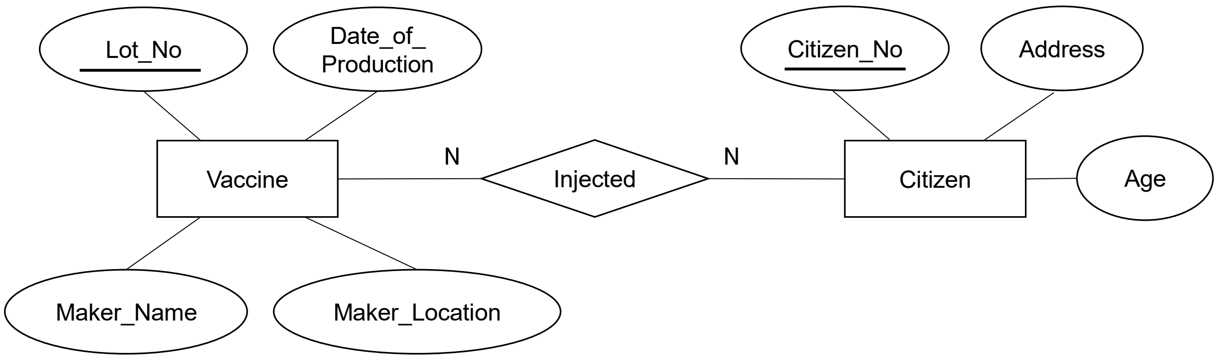


Fig. 1: Entity-Relationship Diagram

[Choices]

(a)	Vaccine	(b)	Citizen	(c)	Injected	(d)	Lot_No
(e)	Citizen_No	(f)	Lot_No, Citizen_No	(g)	Date_of_Vaccination	(h)	Date_of_Production
(j)	Maker_Name	(k)	Maker_Location	(m)	Address	(n)	Age
(p)	first	(q)	second	(r)	third	(s)	Boyce-Codd

Question 2. For the following explanation on database concurrency control, choose a word or phrase that best fits the enclosed parts with symbols (a) through (e), and answer with the number of your choice. The choices and their numbers are described in the enclosing box. If you do not find a suitable choice, write a suitable word or phrase. Assume that the same word or phrase goes in boxes with the same symbol.

[(a) ① Relations ② Schemas ③ Transactions ④ Indexes] represent a collection of indivisible actions, which must satisfy the property called [(b) ① uniformity ② atomicity ③ integrity ④ exclusiveness], where either all actions are executed or none are. For the concurrency control to keep the consistency of data when multiple [(a)] are executed simultaneously, the method called [(c) ① logging ② two-phase commit ③ two-phase locking ④ slicing] is used. In case of using [(c)], [(d) ① unlock ② disk failure ③ bug ④ deadlock] may occur where all [(a)] wait for each other. To release [(d)], it is necessary to terminate [(e) ① all ② one of] the [(a)] in the waiting status.

Question 3. When expressing the following relational algebra operation in a SQL statement, answer the appropriate expressions that should be placed in the numbered brackets.

Relational algebra operation: $\pi_{P.a,Q.d}(\sigma_{Q.d<20}(P \bowtie_{P.b=Q.c} Q))$

SQL Statement: `SELECT [①] FROM [②] WHERE [③] AND [④]`

Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

Computer Science⑨ Artificial Intelligence　　＊This question consists of 2pages.

Answer all the questions below.

Question.　For the following explanations on Artificial Intelligence (AI) from (i) to (v), choose the most suitable word, equation, or number for the boxes (1) to (15) from the options and answer with the symbols shown in letters of the alphabet. Note that some of the unrelated options are included, and some of the letters are omitted. Assume that the same word is enclosed in the box with the same number.

(i)　There are a best first search algorithm and A* algorithm for pathfinding. The best first search expands the nodes with the lowest [(1)] and A* algorithm with the lowest [(2)], respectively. In the following graph, the best first search finds [(3)] as a path from S to G and A* algorithm obtains [(4)]. A* algorithm is guaranteed to find the optimal path if the estimated cost is [(5)] the actual cost, while the best first search does not guarantee the optimality.

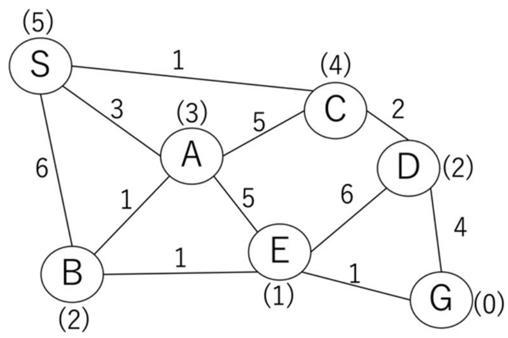


Fig. 1. Graph (The numbers along edges are actual costs and ones in parentheses are estimated costs)

(ii)　The multi-stage decision problem is a problem to maximize the sum of profits when a chosen action at time t decides the state at time t+1 and obtains a profit. The efficient method to solve the multi-stage decision problem by breaking it down into sub-problems is [(6)]. [(6)] applies [(7)] to reuse values that have already been calculated.

(iii)　Bayes' theorem is utilized to predict an event in the uncertain real world. The Bayes' theorem is stated mathematically as the following equation, [(8)], where A and B represent independent events. This theorem is applied to a spam filter. Assume that the following table summarizes each probability of word w_1, w_2, and w_3 occurring for spam and not spam. If spams occur with probability 0.1, the probability of the spam occurring for a mail containing w_2 is [(9)]. Also, the probability of spam for a mail containing w_3 is [(10)] than one of not spam.

	w_1	w_2	w_3
$P(w_i \mid spam)$	0.5	0.1	0.4
$P(w_i \mid not\ spam)$	0.5	0.3	0.2

(iv)　Symbolic logic is a formalization method that expresses logical relationships among events described in languages by converting them into symbols. Especially, the logic that expresses propositions by splitting it into a subject and predicate and using variables is called predicate logic. For example, "all people like peace." is expressed in predicate logic as below.

[(11)] (human(x) [(12)] like(x, PEACE))

Note that human(x) and like(x, y) are predicates that express "x is human" and "x likes y" and PEACE is a constant expressing "peace".

(v)　A supervised learning uses [(13)] containing supervisory signal for learning. This learning method updates parameters to reduce errors between the supervisory signal and output value. The technique to update the parameters based on a gradient of the error function is called [(14)]. This type of learning can be used for [(15)], which outputs a real value given an input value, and classification, which classifies inputs into categories.

Master's Program, Information Science and Engineering Course,
Graduate School of Information Science and Engineering,
Ritsumeikan University

＊This question consists of 2pages.

[Options]

a	$P(B\|A) = \frac{P(A\|B)P(B)}{P(A)}$	b	Gradient Ascent	c	Estimated cost	d	$P(A,B) = P(A\|B)P(B)$
e	Optimization problem	f	Sum of actual cost and estimated cost	g	$P(A) = \sum_{B} P(A,B)$	h	∃x
j	More than	k	Training data	l	Linear programming	m	Maximization
n	∀x	o	Memoization	p	Reward	q	→
r	Actual cost	s	Higher	t	∧	u	Lower
v	Dynamic programming	w	Test data	x	Less than	y	Gradient descent
z	Regression	A	Minimax	B	Hashing	D	Search problem
E	27/100	G	1/100	H	1/28	J	27/28
L	S->B->E->G	M	S->A->B->E->G	Q	S->C->D->G	R	S->C->E->G

Specialized Subjects

Human Information Science

⑩Image Processing

⑪Artificial Intelligence

In case choosing the Human Information Science section, choose one question either ⑩ or ⑪

This page is blank

Master’s Program, Information Science and Engineering Course,
Graduate School of Information Science and Engineering,
Ritsumeikan University

Human Information Science⑩ Image Processing ＊This question consists of 4 pages.

Answer all the questions below.

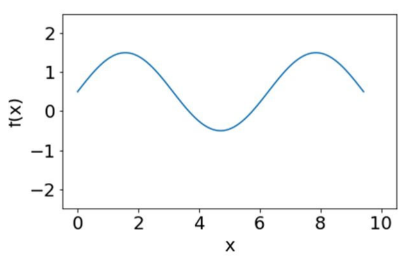
Question 1. Answer each question.

(1) The analog video signal standard “NTSC”, once employed in television broadcasting in Japan, adopted YIQ color coordinate system for its color representation. A color (R,G,B) represented in RGB color coordinate is converted to (Y,I,Q) in YIQ color coordinate by the following equations:

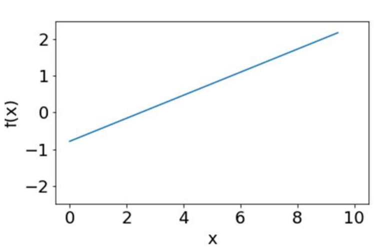
$$X_1 = 0.596R - 0.275G - 0.321B,$$
$$X_2 = 0.212R - 0.523G + 0.311B,$$
$$X_3 = 0.299R + 0.587G + 0.114B.$$

Of $X_1, X_2, and\ X_3$, which corresponds to the signal Y of YIQ color coordinate that represents the brightness (intensity) of pixels? Answer with the reason for your choice.

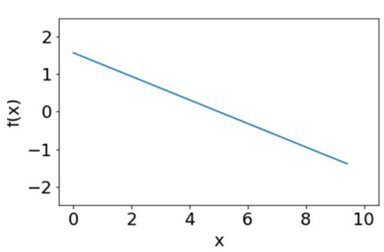
(2) SSD (Sum of Squared Difference) and correlation coefficient are often used for solving stereo feature correspondences by pixel pattern matching. Choose the most appropriate signal profile that gives the largest correlation coefficient with the 1-D signal $f(x)$ shown below among the following choices (a) to (d).



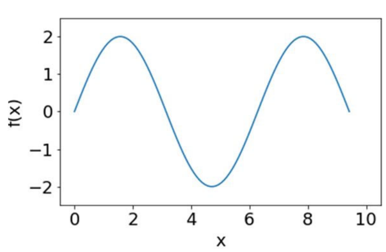
Signal $f(x)$



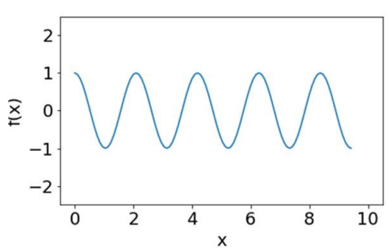
(a)



(b)



(c)



(d)

Choices

＊This question consists of 4 pages.

Question 2. For a grayscale image sequence $I(x, y, t)$ which captures a moving object, where (x, y) is the horizontal and vertical image coordinates and t is time, we want to obtain the optical flow field $\left(\frac{dx}{dt}, \frac{dy}{dt}\right) = \left(u(x, y, t), v(x, y, t)\right)$, the motion information for each pixel, by gradient method. The spatio-temporal gradient constraint equation for u and v is given as following:

$$\frac{\partial I(x, y, t)}{\partial x} u(x, y, t) + \frac{\partial I(x, y, t)}{\partial y} v(x, y, t) + \frac{\partial I(x, y, t)}{\partial t} = 0.$$

(1) Let $I(x, y, t)$ be the pixel value at (x, y) of the image captured at time t. When the optical flow at pixel (x, y) is assumed to be $\left(u(x, y, t), v(x, y, t)\right)$, then the point on the object projected onto this pixel should be projected onto the pixel $(x + u\Delta t, y + v\Delta t)$ in the image taken after small time duration Δt, and its pixel value should be $I(x + u\Delta t, y + v\Delta t, t + \Delta t)$. Assuming that these two pixel values are equivalent, derive the spatio-temporal gradient constraint equation above. (Hint: Employ the first-order approximation of Taylor expansion.)

(2) Only one spatio-temporal gradient constraint equation can be obtained for each pixel and it is not enough to uniquely determine (u, v) for every pixel. This is called *the aperture problem*. How can we obtain the optical flow $\left(u(x, y, t), v(x, y, t)\right)$ given the existence of the aperture problem? Show and explain an effective solution in detail.

＊This question consists of 4 pages.

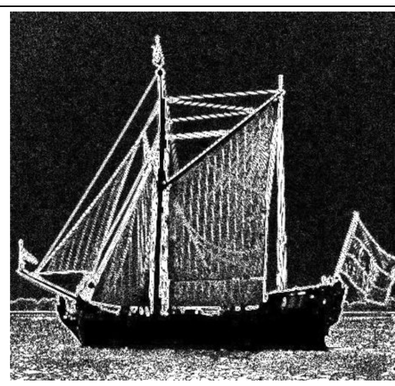
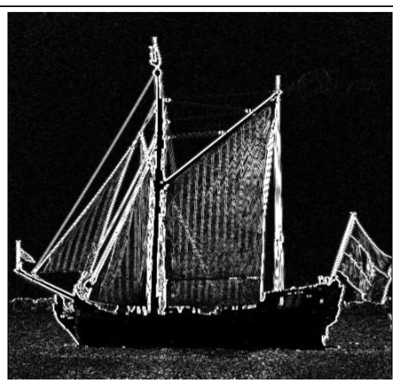
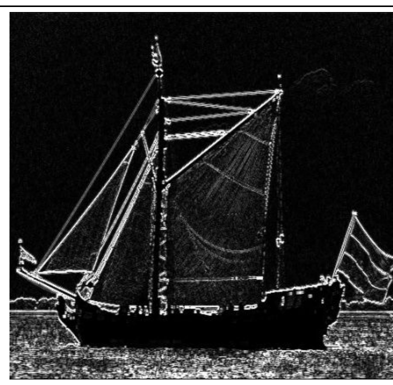
Question 3. In image understanding by using deep neural networks, "Conv layer" based on the convolutional filter is often employed as the feature extraction layer. Choose the most appropriate image for the output obtained by applying each of 3x3 convolutional filters (1) ,(2), and (3) to the input image below.

Input image (grayscale)

Convolutional filters

(1)			(2)			(3)		
-1	0	1	-1	-1	-1	-1	-2	-1
-2	0	2	-1	8	-1	0	0	0
-1	0	1	-1	-1	-1	1	2	1

Choices of output images

(a) 	(b) 	(c) 
(d) 	(e) 	(f) 

＊This question consists of 4 pages.

Question 4. In video streaming services via computer network, compression of image data allows live streaming in realtime even if network has a limited transfer bitrate. When the average transfer bitrate between the streaming server and user’s terminal is 20 Mbps (Mega bits per second), what percentage must the amount of compression be compared to the original one for each video frame in order to provide video streaming at 30 fps (frame per second) in average?
Each original video frame has a resolution of 1000x1000 pixels, consists of three planes of RGB color, and the pixel value of each plane is expressed in 8-bit depth (i.e., 24 bits in total for one pixel). You may assume an ideal case where the fluctuation of network speed is extremely small.

Master's Program, Information Science and Engineering Course, Graduate School of Information Science and Engineering, Ritsumeikan University

Human Information Science⑪ Artificial Intelligence

Question 1. Briefly explain all the phrases below with a specific example.

(1) Constructive approach

(2) Symbol grounding problem

(3) Nash equilibrium

(4) Markov decision process

Question 2. Select the most suitable phrase to fill in the blanks in the following sentence from the choices below (write the number).

Machine learning is classified into reinforcement learning, supervised learning, and unsupervised learning, depending on how feedback information is handled. Unsupervised learning is generally used for clustering and dimension reduction of data. As a typical clustering method, the [(1)] is known, and as a dimension reduction method, [(2)] is known. Problems tackled by machine learning are roughly divided into [(3)] problems and classification problems. In the classification problem, the naive Bayes classifier is a [(4)] model, and the support vector machine is a [(5)] model.

【Choices】①Bayes' theorem ②fuzzy theory ③differential ④principal component analysis ⑤k-means algorithm ⑥Integral ⑦multinomial distribution ⑧generative ⑨Gaussian distribution ⑩imitation ⑪transfer ⑫discriminative ⑬normal distribution ⑭regression ⑮beta distribution

Question 3. Select the most suitable phrase to fill in the blanks in the following sentence from the choices below (write the number).

The basic formula of probability is utilized in the construction of artificial intelligence. It is called the multiplication theorem that a joint probability P (A, B) is rewritten as P (A, B) = P (A | B) P (B). In this equation, P (A | B) is called [(1)]. If P (A, B) is marginalized for B, it becomes [(2)]. Bayes' theorem is an equation derived from the [(3)]. The posterior probability, which is the probability of the cause for the resulting event, can be obtained from likelihood and the prior probability. The graphical model shows the [(4)] between random variables. A set of nodes [(5)] defined by P (A | ∂A, B) = P (A | ∂A) is called a Markov blanket.

【Choices】①multiplication theorem ②P(B|A) ③addition theorem ④conditional probability ⑤joint probability ⑥posterior probability ⑦dependencies ⑧P(B) ⑨expected value ⑩neural network ⑫∂A ⑬A ⑭conditional expectation ⑮P(A)

Question 4. Convert the following propositional logical formula to the conjunctive normal form.

$P \equiv Q \vee R$

This page is blank