February 2021 Admissions

AY2021 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.

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

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.

Common Subjects	①Linear Algebra ②Probability and Statistics ③Data Structure and Algorithms	
Specialized Subjects	Computer Science	④Computer Architecture ⑤Operating System ⑥Software Engineering ⑦Computer Networks ⑧Databases ⑨Artificial Intelligence
	Human Information Science	⑩Image Processing ⑪Artificial Intelligence

【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

Answer all the questions below.

Question 1. Find the determinant $|X|$ of the $(N+1)\times(N+1)$ matrix X. Note that a is a constant and $x_i = i$.

$$X = \begin{pmatrix} a & 0 & \cdots & 0 & x_1 \\ -a & a & \ddots & \vdots & \vdots \\ 0 & -a & \ddots & 0 & x_{N-1} \\ \vdots & \ddots & \ddots & a & x_N \\ 0 & \cdots & 0 & -a & a \end{pmatrix}$$

Question 2. Find the orthonormal basis for the following set of vectors.

$$\left\{ \boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Question 3. Given two matrices $Y = \begin{pmatrix} 5 & -3 & 6 \\ 2 & 0 & 6 \\ -4 & 4 & -1 \end{pmatrix}$ and $Z = \begin{pmatrix} 4 & -3 & 6 \\ 2 & -1 & 6 \\ -4 & 4 & -2 \end{pmatrix}$, find Y^n and Z^n.

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

Common Subjects ② Probability and Statistics

Answer all the questions below.

Question 1. Find the constants C_i $(i = 1,2,\dots 5)$, the expectation, and the variance of the following probability distribution $P(X = k)$ for discrete variable X, or probability density distribution $f(x)$ for continuous variable x. Answer the questions with the derivation process.

（1） $P(X = k) = \begin{cases} C_1 \cdot k, & (k = 1,2,3,4), \\ 0, & (\text{Otherwise}) \end{cases}$

（2） $f(x) = \begin{cases} C_2 \cdot x, & (0 < x \leq 4), \\ 0, & (\text{Otherwise}) \end{cases}$

（3） $P(X = k) = \frac{C_3}{k!(100-k)!} 0.1^k 0.9^{100-k}, \quad (k = 0,1,2,\dots,100)$

（4） $f(x) = \begin{cases} C_4 \cdot \exp(-2x), & (x > 0), \\ 0, & (\text{Otherwise}) \end{cases}$

（5） $f(x) = C_5 \cdot \exp(-x^2 + 2x), \quad (x \in \mathbf{R})$

Question 2. Answer the following questions with the derivation process.

（1） In a certain examination, each examinee had to choose between two questions A and B. As a result of scoring, the score X of question A follows the normal distribution $N(60,\ 8^2)$, and the score Y of question B follows the normal distribution $N(55,\ 10^2)$. Therefore, the score X of question A is adjusted by a linear transformation to make the two distributions match. Find the appropriate formula of the linear transformation.

（2） The infection rate of a certain infectious disease becomes 1%. Therefore, it is decided to check each individual for the infection by a certain primary test. It has been found for this primary test that 99% of infected individuals are correctly positive, while the remaining 1% are falsely negative. In addition, 90% of uninfected individuals are correctly negative, while the remaining 10% are falsely positive. Now, find the probability that a person who is positive in this primary test is truly infected. Give the probability in an irreducible fraction.

Question 3. Describe the following for two random variables X and Y.

（1） Definition of the independence of X and Y.

（2） Definition of (Pearson's) correlation coefficient between X and Y.

（3） Derivation of the proof that if X and Y are independent then they are uncorrelated.

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. Given the number of elements in the input array as n, classify bubble sort, heap sort, merge sort, bucket sort, selection sort, and insertion sort to A), B), or C), respectively.

A) Non-comparison sort

B) Comparison sort, worst-case time complexity is $O(n^2)$

C) Comparison sort, worst-case time complexity is O(n log n)

Question 2. Figure 1 is pseudo-code for quick sort. X[p] represents the p-th element of the array X. *Partition* is a function that selects the last element of the interval from index p to r in array X as a pivot, and relocates each element smaller than the pivot to the first half of the array and each larger element to the second half of the array. *Swap* is a function that exchanges two values. Answer the following questions.

（1） Answer the pseudo-code corresponding to the blanks of (A) through (D) shown in Figure 1.

（2） Illustrate the process of changing the value of array X when performing the Quicksort(X,0,4) of Figure 1 for array X of Figure 2.

（3） What is the time complexity of the average-case and worst-case of Quicksort, expressed in order notation?

（4） Explain why the worst case occurs in Quicksort and how to improve the algorithm.

```
Quicksort(X, p, r)
   if p < r
      pos = Partition (X, p, r)
      Quicksort (X, p, (A) )
      Quicksort (X, pos+1, (B) )

Partition(X, p, r)
   pivot = (C)
   i = p-1
   for j = p to r-1
       if X[j] < (D)
          i = i+1
          Swap(X[i], X[j])
   Swap(X[i+1], X[r])
   return i+1
```

Figure 1: Pseudo-code of quick sort

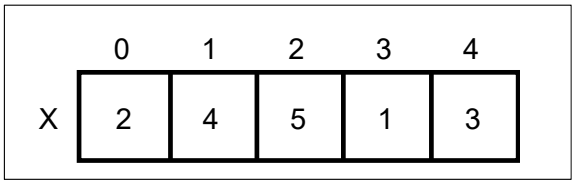


Figure 2: Array X

Question 3. Let every node of a binary search tree have a unique value (i.e. no duplication). Answer the following questions.

（1） Answer which of the following A) or B) is the appropriate rule for the nodes of binary search tree.

A) All values of the nodes on the left subtree of a node v are less than its own node value, and all values of the nodes on the right subtree of a node v are greater than its own node value.

B) The value of each parent node is greater than the values of its child nodes.

（2） Illustrate the binary search tree created when {10, 12, 5, 4, 20, 8, 7} are inserted into the binary tree in this order.

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 details of implementation and the advantages compared to others for each of these three cache placement policies: direct-mapped, fully-associative, and set-associative, while making each difference clear. You may use figures in the explanation.

Question 2.

Explain how and when loss of significance on floating point operations occurs using an example. You may suppose IEEE754, another floating point representation, or a representation defined in your explanation.

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

Computer Science ⑤ Operating System

This is a question about page replacement in virtual memory management of an operating system. Suppose that 6 pages which are numbered from 0 to 5 in virtual address space are in use. Also, suppose that there are 3 page frames in real memory. When the pages are referenced as shown in the following tables, which use FIFO or LRU as a page replacement algorithm, write "X" for page fault occurrence due to reference, and write the content of the page frames. The content of page frame must be written as a queue for FIFO, or as a stack for LRU.

FIFO

Page Fault																
Reference String	0	1	2	3	1	4	1	5	2	3	5	0	5	3	1	3
Page frames (FIFO Queue)																

LRU

Page Fault																
Reference String	0	1	2	3	1	4	1	5	2	3	5	0	5	3	1	3
Page frames (LRU Stack)																

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

Computer Science ⑥ Software Engineering

Answer all the questions below.

Question 1.
Describe the differences between requirements and requirements specification.

Question 2.
Explain the client-server model in architectural design.

Question 3.
Describe situations where adaptive software maintenance is conducted.

Question 4.
Explain the relationship between classes and instances in object-oriented software development.

Question 5.
Explain the pros and cons of using the waterfall model for software development.

Question 6.
Explain the role of stubs in top-down integration testing.

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

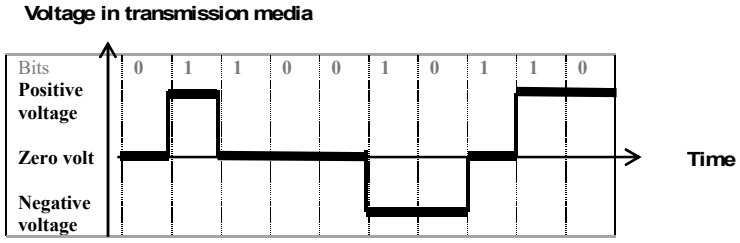
Computer Science ⑦ Computer Networks

Answer all the questions below.

Question 1. Answer the following two sub-questions regarding line codes that use positive and negative voltages as physical quantities.

(1) Describe these two types of line code: (Q1a) a non-return-to-zero level (NRZ(L)) line code and (Q1b) Manchester code (or phase encoding).

(2) Illustrate changes in physical quantities of transmission media such as voltage for bit sequence "0110010110" for each of the two line codes: (Q2a) non-return-to-zero line code and (Q2b) Manchester code. The following is an example of another line code to show how to indicate changes of voltage, for reference.



Example showing how to illustrate changes of voltage for MLT-3 (Multi-Level Transmit).
The line code is the bold line part. The vertical dotted lines indicate the transmission clock rise timings.

Question 2. For the descriptions regarding computer networks from (1) to (7), complete each description by selecting correct answers from options inside the rectangles. You can answer by writing suitable keywords or phrases if you find no suitable options. Assume that the same option should be assigned to the rectangles with the same question identification in the description.

(1) The acronym of an industry-driven interest group that was founded in 1993 that provides specifications for a complete set of protocols for wireless infrared communications is [Q(a) ① ITU ② IETF ③ IEEE ④ IrDA] and the name [Q(a)] also refers to that set of protocols.

(2) In the [Q(b) ① ALOHA ② CSMA ③ FDDI ④ IEEE 802.5] protocol, which is a bus or media access control algorithm, each node does not sense whether there is another node's signal on the bus or not, and starts to transmit packets as soon as they are ready.

(3) In the Open Systems Interconnection model, or OSI model, the [Q(c) ① physical ② data link ③ network ④ transport] layer is responsible for packet forwarding, including routing through intermediate routers and networks. In a LAN, a device that relays frames in the middle of a transmission path to extend the transmission distance and only transfers them at the data link layer of the OSI model is called [Q(d) ① a repeater ② a router ③ a network bridge ④ an ONU] .

(4) A host with IPv4 (Internet Protocol version 4) address in dotted decimal notation (or dot-decimal notation) 192.168.25.254 belongs to class [Q(e) ① A ② B ③ C ④ D ⑤ E] networks. If we change the subnet mask of the network to 255.255.255.192 in dot-decimal notation, the maximum number of unique host IP addresses is [Q(f) ① 62 ② 168 ③ 192 ④ 254] .

(5) The first four bits of the IP header indicate the version number of the IP, and a binary [Q(g) ① 0001 ② 0100 ③ 0110 ④ 1001] indicates that IPv6 (Internet Protocol version 6) is being used. The Internet Header Length field, which indicates the size of the protocol header, is [Q(h) ① in IPv4 header only ② in IPv6 header only ③ in both IPv4 and IPv6 headers] .

(6) Regarding TCP and UDP on IPv4, [Q(i) ① only TCP has ② only UDP has ③ both TCP and UDP have] the source port number field on each header.

(7) The [Q(j) ① RIP ② OSPF ③ BGP-4 ④ CIDR] is a dynamic routing protocol that automatically determines the shortest path using the number of routers traversed as the distance metric. It is difficult to apply it in large networks for reasons such as the inability to inform neighboring routers of the subnet mask information.

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

Computer Science ⑧ Databases

Answer all the questions below.

Question 1. Complete the 3 sentences ((1)-(3)) about the database by selecting a number from each box (A.- J.). In the case of a blank box or no answer from a box, write a word for the answer. Assume that the same answer is assigned to each box with the same symbol.

（1） [A.] is an indivisible series of database operations. The properties of a reliable [A.] are defined as [B.] criteria. [C. ① Consistency ② Durability ③ Atomicity ④ Isolation] means all of the [A.] are completed or none are completed.
[D. ① Consistency ② Durability ③ Atomicity ④ Isolation] represents that the results of executing [A.] concurrently are the same as ones of executing them sequentially. The execution plan of [A.] that satisfies [D.] is called [E. ① parallelizable ② concurrentable ③ serializable ④ encodable]. The execution plan in Figure 1 [F. ① is ② is not] [E.]. Note that Q1, Q2, Q3, and Q4 in Figure 1 indicate [A.], and Read and Write are operations to read and write data of the argument from and to the database, respectively.

Time	Q1	Q2	Q3	Q4
t_1	Read(X)			
t_2		Read(X)		
t_3	Write(X)			
t_4		Read(Y)		Read(X)
t_5			Read(X)	
t_6		Write(Y)		
t_7				Read(Y)
t_8			Write(X)	Write(Y)

Figure 1: Execution plan

（2） A relation is described as *Name*(*Attribute*, …) and the attribute of a primary key is underlined. The Relation *Order*(*Order No.*, *Product No.*, *number*, *unit price*, *total price*) is the [G.] normal form.

（3） B tree is often used to make an index for a database. After inserting 26 to the B tree illustrated in Figure 2, a, b, and c, of the updated B tree in Figure 3 are [H.], [I.], and [J.], respectively.

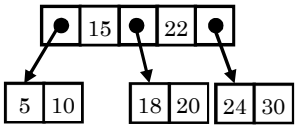


Figure 2. B tree

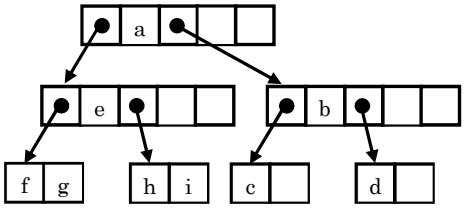


Figure 3. B tree after inserting 26

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

Computer Science ⑨ Artificial Intelligence ＊This question consists of 2pages.

Answer all the questions below.

For the following explanations on Artificial Intelligence (AI) from (i) to (iv), choose the most suitable word for the boxes (1) to (15) from the options and answer with the symbols shown by the alphabetic letters. Note that some unrelated options are included, and some of the letters are omitted. Assume that some words should be selected for boxes with the same numbers.

(i). A* algorithm, one of the methods for efficient pathfinding, is executed using the sum of the actual cost from the initial node to the current node and the estimated cost from the current node to the goal node. The search is proceeded by sorting the elements in the [1], which is a set of candidate nodes. In order to secure the optimality of the obtained path, it is required for the estimated cost from the current node to the goal node to be equal to or [2] than the actual cost.

(ii). The operation of classifying a collection of data into several groups based on the similarity between the data is called [3]. [4] is the most basic [3] method and is based on an algorithm that is easy to understand intuitively, that a set of n-dimensional feature vectors can be classified into K clusters. In the algorithm, first, the data is divided into initial clusters. Then, each data is attributed to the closest cluster based on its distance from the cluster's [5]. Finally, repeat the process of updating each [5] until the cluster assignments are no longer changed. While it is easy to understand, the [4] has a disadvantage that it cannot properly classify data whose shape of data distribution differs for each cluster. In order to classify the data into clusters with different coverage in the feature space, there is a [3] approach that relies on a [6].

(iii). Reinforcement learning is a method of machine learning that allows an agent (e.g., a robot) to learn through trial and error, and is a method of learning behavior that maximizes the sum of [7], which can be obtained in the future. Reinforcement learning is formalized based on [8], and its goal is to acquire better policies. In order to learn better policies, we need to appropriately estimate the states in which the agent is placed and the value of its behaviors in each state. Therefore [9] is defined.

(iv). Symbolic logic is the study of how to understand things expressed in languages based on their logical relationships after converting them into symbols, and the most basic one is propositional logic. [10] is an extension of propositional logic by dividing propositions into subjects and predicates. The following five logic symbols are used in [10]:

- ¬: negation
- ∧: [11]
- ∨: [12]
- →: implication
- ≡: equivalence

The relationship between the Boolean values of the logical formulas P and Q and the Boolean values of the logical formulas combined with the above logic symbols can be shown as follows

P	Q	¬P	P∨Q	P∧Q	P→Q	P≡Q
1	1	0	1	1	1	1
1	0	0	1	0	0	0
0	1	1	1	0	[13]	0
0	0	1	0	0	1	1

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

---

When a logical formula is always true, regardless of any interpretation, the formula is called a [14]. The following is one of the typical [14], known as [15].

$$\neg\ (P \wedge Q) \equiv \neg P \vee \neg Q,\quad \neg(P \vee Q) \equiv \neg P \wedge \neg Q$$

a	Less	b	Markov Decision Process(MDP)	c	Contradiction	d	Discount
e	Tautology	f	Estimation function	g	Clustering	h	Conjunction
j	Representative value	k	Open list	m	1	p	Rewards
q	Disjunction	r	Predicate logic	s	EM algorithm	t	0
u	Closed list	w	Sampling	x	More	y	Value function
z	Ontology	A	Modal logic	B	Predicted value	D	Conjunctive normal form
E	Gaussian Mixture Model	G	K-means clustering	H	de Morgan's law	R	Center

Specialized Subjects

Human Information Science

⑩Image Processing
⑪Artificial Intelligence

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

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

Human Information Science ⑩ Image Processing ＊This question consists of 4 pages.

Answer all the questions below.

Question 1. The following questions are about spatial filtering.

(1) Which spatial filter gives the same effect as filtering an image by applying the filters shown in Figure 1, first A, then B? Choose the most appropriate one from the options.

(A)

0	1	0
0	1	0
0	1	0

(B)

0	0	0
-1	0	1
0	0	0

Figure1

【Options】

(a)

-1	0	1
-1	0	1
-1	0	1

(b)

1	0	-1
1	0	-1
1	0	-1

(c)

0	1	0
1	-4	1
0	1	0

(d)

0	1	0
-1	1	1
0	1	0

(2) The filter expressed in equation ① is a LoG filter which consists of Gaussian filter and Laplacian filter.

$$\mathrm{LoG}(x,y) = \frac{x^2+y^2-2\sigma^2}{2\pi\sigma^6}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots ①$$

When detecting zero crossings from images generated from the LoG filter, the value of σ in equation ① needs to be adjusted for a certain reason. Choose the most appropriate reason from the options below.

【Options】

(a) Faster smoothing and edge detection

(b) Faster zero crossings calculation

(c) Calculating the similarity between gray-level distribution and Gaussian distribution

(d) Adjusting the thickness of detected edges

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

Question 2. Considering a stereo camera system (see Figure 2) in which two cameras are aligned in parallel with interval b, and the x-axes of two camera coordinate systems are on the same line. The origins of camera image plane coordinate systems C_L and C_R are the intersections of the optical axes with image planes. The positions in 3D space are represented in the camera coordinate system with the origin O set to the left camera's optical center. Also, f is the focal length. Answer the following questions and choose the most appropriate answer from the options.

(1) A point $P(X,Y,Z)$ in the 3D space is mapped onto points $P_L(x,y)$ and $P_R(x',y')$ on the left and right image plane coordinate systems, respectively. Indicate how to represent the coordinate of point $P(X,Y,Z)$ in 3D space.

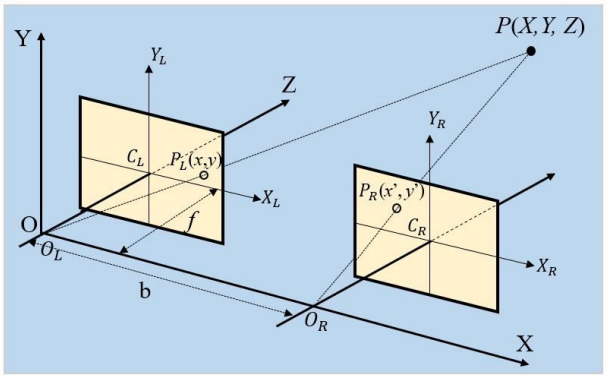


Figure 2

【Options】

(a) $X = \frac{b(x-x')}{x}$ $Y = \frac{b(x-x')}{y}$ $Z = \frac{b(x-x')}{f}$

(b) $X = \frac{x}{b(x-x')}$ $Y = \frac{y}{b(x-x')}$ $Z = \frac{f}{b(x-x')}$

(c) $X = \frac{bx}{x-x'}$ $Y = \frac{by}{x-x'}$ $Z = \frac{bf}{x-x'}$

(d) $X = \frac{x-x'}{bx}$ $Y = \frac{x-x'}{by}$ $Z = \frac{x-x'}{bf}$

(2) Given the point $P_L(x,y)$ in the left image, select the most appropriate description of its epipolar line in the right image.

【Options】

(a) Parallel to X_R axis

(b) Parallel to Y_R axis

(c) Parallel to the line passing through the origin C_L and the point $P_L(x,y)$

(d) Perpendicular to the line passing through the origin C_L and the point $P_L(x,y)$

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

---

Question 3. The following questions are about image restoration. The process of image distortion due to defocusing and blurring can be expressed by equation ② using point spread function $h(x, y)$. Here, the original image before distortion is represented by $f(x, y)$ and the distorted image is represented by $g(x, y)$, where x, y are coordinates in the image.

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-\xi, y-\eta)h(\xi,\eta)d\xi d\eta \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots ②$$

（1） $F(u, v)$ and $H(u, v)$ are the two-dimensional Fourier transform of $f(x, y)$ and $h(x, y)$. The convolution integral is represented by $*$. Assuming the two-dimensional Fourier transform of $g(x, y)$ is $G(u, v)$, choose the most appropriate representation of $G(u, v)$ from the options below.

【Options】

(a) $G(u, v) = F(u, v) * H(u, v)$

(b) $G(u, v) = F(u, v)H(u, v)$

(c) $G(u, v) = \frac{H(u,v)}{F(u,v)}$

(d) $G(u, v) = \frac{F(u,v)}{H(u,v)}$

（2） Gaussian distribution is one of the functions for approximating the point spread function of image defocusing. Using Gaussian distribution, the point spread function $h(x、y)$ can be approximated as equation ③, where σ is the parameter to represent the extent of blurring. How is the two-dimensional Fourier transform of the point spread function represented? Choose the most appropriate representation from the options below.

$$h(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots ③$$

【Options】

(a) Bessel function of the first kind of order two

(b) Bessel function of the first kind of order one

(c) Gaussian distribution

(d) Laplace distribution

Question 4. The following questions are about pattern recognition. Choose the most appropriate answer from options for questions (1)～(2)

（1） Features are extracted from images belonging to class 1 or class 2 and plotted in a two-dimensional feature space. The distribution of features is shown in Figure 3 (● represents class 1, and ▲ represents class 2). These features are used to classify three new test images (images ①～③, represented by × in the feature space). Select the option that consists of all images that would be classified into different classes when using these two algorithms: nearest neighbor method and k-nearest neighbor method with k=3, measured by Euclidean distance.

【Options】

(a) ①
(b) ②
(c) ① and ②
(d) ③
(e) ① and ③
(f) ② and ③

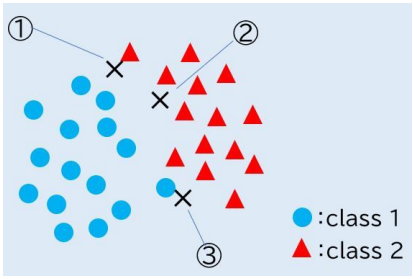
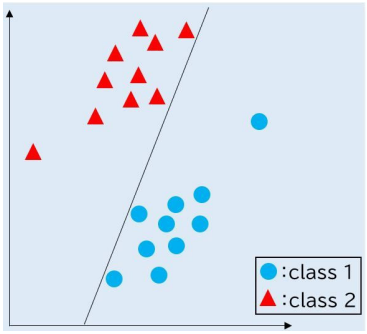


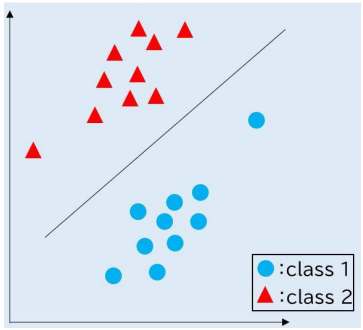
Figure 3

（2） A support vector machine (SVM) is a discriminator for making the decision boundary between two classes. Which of the following results in Figure 4 represents the decision hyperplane separating two classes of training samples using a hard margin linear SVM?

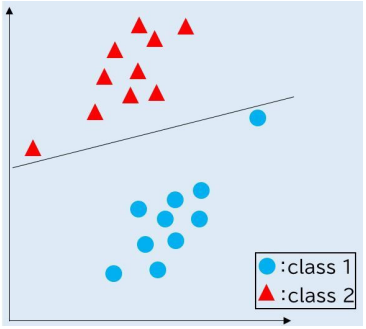
【Options】



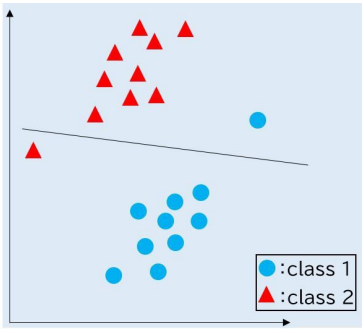
(a)



(b)



(c)



(d)

Figure 4

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

Human Information Science ⑪ Artificial Intelligence

Answer all the questions below.

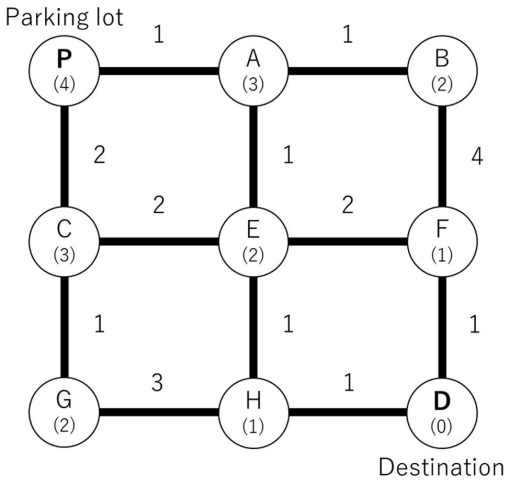
Question 1. An autonomous vehicle needs to find the shortest path from a parking lot (**P**) to a destination (**D**). The value shown beside each edge means the actual cost of that edge. That inside each node represents the expected future cost to reach the destination. The total cost of travel is the sum of the actual cost on the path from the parking lot to the destination. The problem is to find the path that requires minimal cost.

(1) Execute an **optimal search** (Dijkstra's algorithm) to find the shortest path. An open list and a closed list should be described.

(2) Execute **A* algorithm** and find a solution. An open list and a closed list should be described.

(3) Is it possible to make the solutions of the optimal search and A* algorithm different by changing the expected future cost of **H**? Please explain if it is possible by mentioning the theoretical background.

(4) In a real-world application scenario, how can we determine the cost of an edge? Please explain your idea with an example case.



Question 2.

(1) Please explain what **minimax** for an **extensive-form game** means in game theory with an example.

(2) Please explain what the **naïve Bayes model** means. The generative process of the naïve Bayes model should be described.

(3) Please explain what **Gaussian mixture models** are.

(4) Please explain the relationship between **Bayes filter** and **particle filter** in state estimation (or self-localization).

(5) Please explain **Bag-of-Words representation** in natural language processing.