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.

Common Subjects ① Linear Algebra

Answer all the questions below.

Question 1.

Given the following matrices A and B, find the x, y and z that satisfy $|A| = |B|$. Note that a, b, and c are constants.

$$A = \begin{pmatrix} 2 & 4 & 6 & 8 & 10 \\ 20 & 18 & 16 & 14 & 12 \\ 22 & 24 & 26 & 28 & 30 \\ 40 & 38 & 36 & 34 & 32 \\ 42 & 44 & 46 & 48 & 50 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ x & a & a & a \\ x & y & b & b \\ x & y & z & c \end{pmatrix}$$

Question 2.

Given a matrix $C = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}$, answer the following.

（1） Find the eigenvalues and eigenvectors of the matrix C.

（2） Diagonalize the matrix C.

（3） Find C^n.

Common Subjects ② Probability and Statistics

Answer all the questions below.

Question 1.　Independent events

There are 250 applicants for an examination. Each applicant will be absent with probability 0.8%, independently. Answer the following questions, with the derivation process for (2) and (4).

（1）　Let X be the number of absentees. What is the probability distribution $P(X = k)$ of X? Answer what is the distribution called, the formula of $P(X = k)$ as a function of k, and the range of k.

（2）　Find the expectation and the variance of X.

（3）　Distribution $P(X = k)$ in the above (1) can be approximated to another probability distribution $\tilde{P}(X = k)$. Answer what is distribution $\tilde{P}(X = k)$ called, the formula of $\tilde{P}(X = k)$ as a function of k, and the range of k.

（4）　There are only 248 seats in the examination hall. Find the probability that the number of attendees will <u>not exceed</u> the number of seats. Use the approximate distribution $\tilde{P}(X = k)$ which is considered in the above (3) for the calculation. Furthermore, obtain the approximate probability in an irreducible fraction, by approximating the Napier number (the base of natural logarithm) e as 3.

Question 2.　Bayesian theorem

There are 3 boxes A, B and C. Each box contains 100 balls of red and white in total, and cannot be distinguished from the outside. Box A contains 80 red balls and 20 white balls, box B contains 60 red balls and 40 white balls, and box C contains 20 red balls and 80 white balls. You will randomly choose 1 box among 3 boxes, then draw one ball from the chosen box. Let us find the probabilities that the chosen box is A, B and C, from the color of the drawn ball.

Let A, B, and C denote the events of choosing box A, B, and C, respectively. And let R and W denote the events of drawing a red ball and a white ball, respectively. Answer the following questions with the derivation process, and give the probability in an irreducible fraction.

（1）　Find the conditional probabilities $p(R|A), p(R|B),$ and $p(R|C)$.

（2）　Find the conditional probabilities $p(A|R), p(B|R),$ and $p(C|R)$.

（3）　The color of the ball, which you drew from the chosen box, was red. Then, you put it back in the original box, and again you drew one ball at random from the same box. The second ball was also red. Now, find the probabilities that the chosen box is A, B and C.

（4）　Similar to the above (3), you drew one ball from the chosen box and record its color, and then you put it back to the original box. Again, you drew one ball at random from the same box, and record its color. After the second recording, you come to know that both colors of the first and the second balls are red. Now, find the probabilities that the chosen box is A, B and C.

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

Common Subjects ③ Data Structure and Algorithms

Answer all the questions below.

Problem 1. Consider the directed graph in Figure 1. In this figure, the circles represent vertices, and the numbers represent vertex numbers. The connecting lines with arrowheads are directed edges. Now answer the following questions.
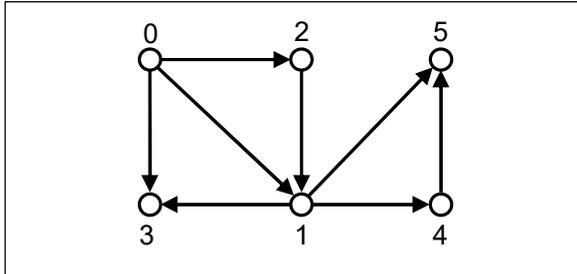

Figure 1: A directed graph

（1） Show the adjacency matrix of the directed graph in Figure 1. Order the adjacent vertices of each vertex according to increasing vertex number.

（2） Using the adjacency matrix derived in (1), starting at vertex 0, perform a depth-first search to visit all the vertices, without repetition. List the vertex numbers in the order they are visited.

Problem 2. Consider a heap represented as a binary tree, as shown in Figure 2. In Figure 2, the circles represent nodes, and the number within a circle represents the data assigned to that node. In order to represent the heap from Figure 2 as an array A, let A[i] represent the i-th index position of the array. Store the root in A[0], the left child of node A[i] in A[2×i+1], and the right child in A[2×i+2]. Then, the pseudo-code to insert a new element (node) into the heap is shown in Figure 3. *size* is a variable that contains the n-1-th index of the array into which n elements have been inserted. $\lfloor x \rfloor$ represents the largest integer less than x. *swap* is a function that exchanges two nodes. Answer the following questions.
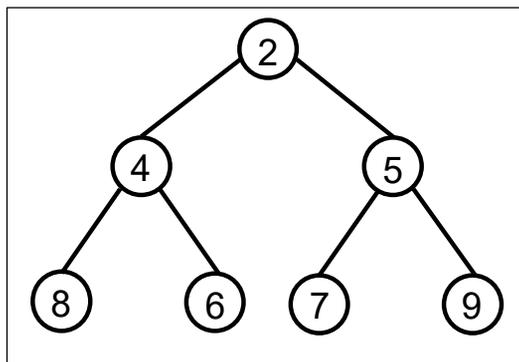

Figure 2: A heap

```
1: insert(key):
2:     size ← size + 1
3:     A[size] ← key
4:     shift-up(size)
5:
6: shift-up(i):
7:     j ← ⌊i/2⌋
8:     if i > 0 and A[i] < A[j] then
9:         swap(A[i],A[j])
10:        shift-up(j)
```

Figure 3: Pseudo-code to insert an element into a heap

（1） Represent the heap in Figure 2 as an array. Also show the indices.

（2） Using the pseudo-code shown in Figure 3, when the function call insert(3) is executed, how many times will the shift-up function be executed? Show the content of vector A at each invocation.

（3） The definition of the shift-up function in Figure 3 calls itself. What is this pattern of invocation called?

（4） Rewrite the shift-up function in Figure 3, so as not to use this pattern (not to call itself), using a *while* statement instead. Show this in pseudo-code, as in Figure 3. There's no need to show line numbers.

（5） If the number of elements stored in the heap is n, what is the worst-case time complexity, expressed in Big-O notation? Concisely explain your reasoning.

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 Science ④ Computer Architecture

Answer all the questions below. Suppose that the target computer system has neither "virtual memory" nor "pipeline control".

(1) Consider the process where a "load" instruction which is one-word, i.e. four bytes long on the main memory at the address A, is fetched and is executed. Explain how the values held in the registers described below are updated in the updating order; program counter, instruction register, memory address register (MAR), memory data register (MDR), and the general purpose register #0 which is also 4 bytes long. When answering describe the "load" instruction at the address A as "X". The instruction X is the instruction to load one-word data Y on the main memory at address B to general purpose register #0.

(2) When considering the three kinds of memory addressing mode; ① immediate addressing, ② direct addressing and ③ indirect addressing, describe how these three addressing modes access the main memory, making the difference of the three modes clear.

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

Computer Science ⑤ Operating System

Answer all the questions below.

(1) About preemptive scheduling algorithms and non-preemptive scheduling algorithms, compare and describe their definitions, mechanisms, and features.

(2) For what kind of systems are each of the algorithm types appropriate? Give examples of those systems, and describe reasons why they are appropriate.

**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.

Explain the meaning of the following technical terms.
 (1) Stakeholder in a software development project
 (2) Non-functional requirement in requirements engineering
 (3) White-box testing for software modules
 (4) Association in object-oriented analysis

Computer Science ⑦ Computer Networks

Answer all problems below.
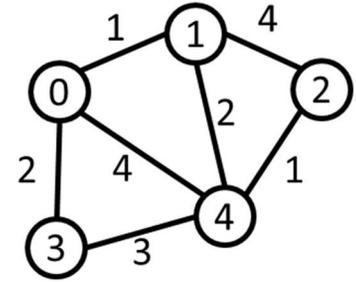
Q1. The figure in the right side shows a network. The number inside the circle denotes a node and its name. The line between two nodes denotes a link between the nodes. The number near the link denotes the distance (or the cost to pass) for the link.

(1) Describe Dijkstra's algorithm by demonstrating all procedures on the network for determining all the shortest distances from node zero "0" to other nodes. Skipping any step or omitting the explanation of each step is not allowed.

(2) Show the order of nodes where the shortest distance is determined by Dijkstra's algorithm in the sub-problem (1). The order should be denoted as the list of node names separated by comma ",". For example, if you want to answer that the nodes 0, 5, 6 and, 7 are the order of shortest distance to be determined, answer that the order of determination is 0, 5, 6, 7.

Also show all shortest paths from node zero to other nodes in this setting. The shortest path should be denoted as the list of node names separated by dash "-" proceeded by the destination node number and colon ":" followed by colon and its distance. If you want to answer that the path from node 0 via node 5 and 7 to node 6 is at a distance of 10, denote it as 6:0-5-7-6:10 for example. Note that the shortest distance from node zero to node zero is determined to be zero at the beginning step of Dijkstra's algorithm, so 0:0-0:0 is self-evident. However, you should also clearly state it in your answer without omission.

Q2. For the descriptions regarding computer networks from (1) to (7), complete each description by selecting correct answers from the 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 descriptions.

(1) The acronym of a specialized agency of the United Nations that is responsible for issues that concern information and communication technologies is Q(a) ① ITU ② IETF ③ IEEE ④ IrDA in English.

(2) One standard for data transmission for LAN, FDDI adopts a ring-based Q(b) ① carrier ② segment ③ collision ④ token network where a signal called a Q(b) is passed between nodes to authorize that node to send messages.

(3) The highest layer of an intermediate open system or intermediate system in the Open Systems Interconnection model or OSI model is Q(c) ① physical ② data link ③ network ④ transport layer. The most similar equipment in Internet protocol suite for an intermediate system is an IP Q(d) ① switching hub ② router ③ bridge ④ repeater .

(4) A host with IPv4 (Internet Protocol version 4) address in dotted decimal notation (or dot-decimal notation) 172.31.26.1 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 255.255.255.128 in dot-decimal notation, the recommended broadcast address for this host is Q(f) ① 172.255.255.255 ② 172.31.255.255 ③ 172.31.26.255 ④ 172.31.26.127 .

(5) The address space of IPv6 (Internet Protocol version 6) is extended to Q(g) ① 48 ② 64 ③ 108 ④ 128 bits while the one of IPv4 is 32 bits. The checksum field in an IP header is Q(h) ① on IPv4 only ② on IPv6 only ③ both on IPv4 and on IPv6 .

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

(7) Q(j) ① Return-to-zero ② NRZM ③ Manchester code ④ 4B5B is a line code in which the encoding of each data bit is either transition of transmission medium state from low to high, or that from high to low, for equal time. It was used in Ethernet 10BASE-T for example.

Computer Science ⑧ Databases *This question consists of 2 pages.

Answer all the questions below.

Question 1. Complete the 3 paragraphs ((1)-(3)) about databases with a number from each box (a.- j.). In the case of finding no answer from a box, write a word for the answer. Assume that the same answer is assigned to any box with the same symbol.

（1） In relational algebra, a theoretical framework for relational databases, a unary operation to obtain the components of a relation restricted to the specific attributes is called
a. ① Selection ② Cartesian product ③ Set union ④ Natural join ⑤ Equijoin ⑥ Projection,
and another one to obtain a subset of a relation that satisfies a certain condition on attributes is called b. ① Selection ② Cartesian product ③ Set union ④ Natural join ⑤ Equijoin ⑥ Projection. A binary operator to combine tuples from two relations where the combination condition is that an attribute from one relation is same as one from the other is called
c. ① Selection ② Cartesian product ③ Set union ④ Natural join ⑤ Equijoin ⑥ Projection,
and another one to remove a redundant attribute from the combined tuples is called
d. ① Selection ② Cartesian product ③ Set union ④ Natural join ⑤ Equijoin ⑥ Projection.
Moreover, the other one to obtain all the combinations of tuples from the two relations is called
e. ① Selection ② Cartesian product ③ Set union ④ Natural join ⑤ Equijoin ⑥ Projection.

（2） SQL, a language to query a database, allows users to combine several Select statements to retrieve data satisfying a complex condition. For example, given a table Student (student_no, student_name, GPA, affiliation), the query to obtain all the attributes of students whose GPA is greater than the biggest GPA among students belonging to the affiliation including "information" is described as below.

```
select  f. ① any ② all ③ some ④ *  from Student
where GPA > g. ① any ② all ③ some ④ *  (select GPA from Student
where affiliation  h. ① in ② = ③ like ④ ==  '%information%');
```

（3） B tree is often used to make an index for a database. After adding 32 to the B tree illustrated in Figure 1, the B tree becomes
i. ① a ② b ③ c ④ d . Then, by adding 7 to the updated B tree,
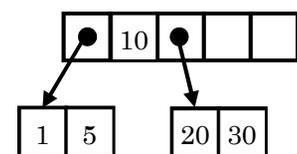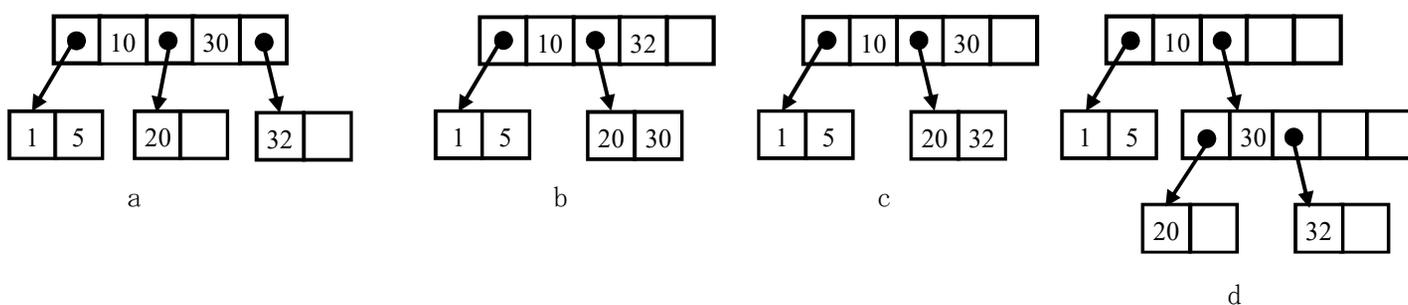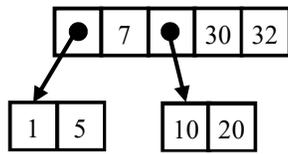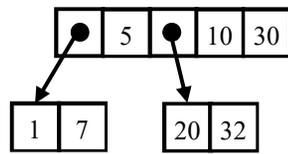it becomes j. ① e ② f ③ g ④ h .
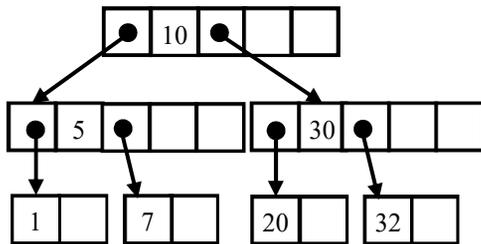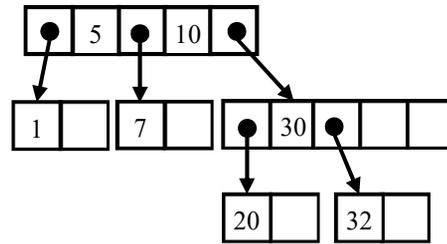

Figure 1: B Tree

a

b

c

d

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

e

f

g

h

Computer Science ⑨ Artificial Intelligence

Answer all the questions below.

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

(i).　A Multi-Agent Systems (MAS) is a system in which multiple agents solve problems in a distributed and cooperative manner. A representative approach to cooperative problem solving by MAS is a (1) protocol that divides a complex task into independent subtasks and assigns them into multiple agents. Among several negotiation protocols, the (2) protocol which models cooperative contract processes in a human society can determine an assignment of subtasks through processes of task announcement, bidding, and contract awarding. (2) has features such as top-down task assignment and (3) between the task manager and contractors.

(ii).　A Neural Network (NN) is an information processing model that mimics human neural systems and can be used for (4) and classification problems. Learning using a NN can be achieved by the iterative modification of the coupling coefficients between neurons composed of multiple layers. (5) is one of the well-known learning methods based on the consecutive modification of coupling coefficients from output layer to input layer for gradually (6) the error between supervisory signal and output value. Recently, (7) , which improves on the shortcomings of (5) , is getting popular due to its higher performance.

(iii).　A* is an algorithm to search a (8) between two nodes on a graph. A* algorithm is executed using the actual cost from the initial node to the current node and the (9) from the current node to the goal node. In order to secure the optimality of the obtained path, it is required for the (9) from the current node to the goal node to be equal to or (10) than the actual cost.

(iv).　Machine learning can be classified to three types such as supervised learning, (11) , and unsupervised learning. (11) is an algorithm to learn an appropriate action based on the (12) for the action.

(v).　Natural Language Processing (NLP) can be achieved with several elementary technologies, such as morphological analysis, (13) , semantic analysis, and context analysis. In the process of morphological analysis, a sentence is divided into morphemes, then (14) , and stemming/lemmatization are conducted. Additionally, for the morphological analysis, word dictionary and linkage rules are used to generate a graph structure ((15)) consisting of candidate words composing a target sentence.

a	Mutual selection	b	Minimized cost	c	Estimated cost	d	Backpropagation
e	More	f	Word lattice	g	Structure analysis	h	Bayesian filter
j	Shortest path	k	Reinforcement learning	m	Reducing	p	Contract net
q	Negotiation	r	Markov Decision Processes	s	Transfer learning	t	Reward
u	Increasing	w	Less	x	Deep learning	y	Longest path
z	Syntactic analysis	A	Segmentation	B	Game tree	D	Particle filter
E	Conjugation	G	Part-of-speech tagging	H	Regression	R	Naive Bayes

Specialized Subjects

Human Information Science

⑩Image Processing
⑪Artificial Intelligence

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

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

Answer all the questions below.

Problem 1.   Questions below are about image sampling and image coding.

(1) Figure 1 shows an 800*800 pixel image with sinusoidal stripe patterns. When resampling the image with equally spaced sampling points in vertical and horizontal grids, in order to satisfy the sampling theorem in both horizontal and vertical directions, the interval between sample points should be smaller than how many pixels? Choose the correct answer from following options.
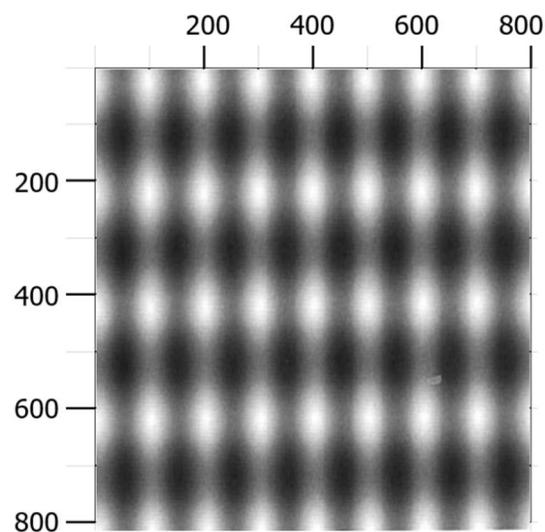
Figure 1

【Options】

a. 25   b. 50   c. 75   d. 100   e. 125

(2) There is an RGB color image with size of 1280*960 pixels and each pixel in R, G, B is quantized to the range 0-255. When the quantization level of R, G, B is set to 32 levels each, without changing the number of pixels in the image, how many bytes will the image be? Choose the correct answer from following options. (Additionally, 1kB=1,024 bytes, 1 byte=8 bits.)
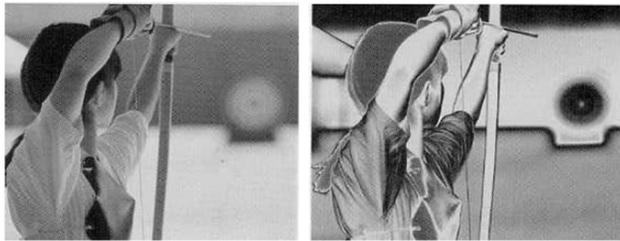
【Options】

a. 576kB   b. 1,728kB   c. 2,250kB   d. 4,608kB   e. 13,824kB

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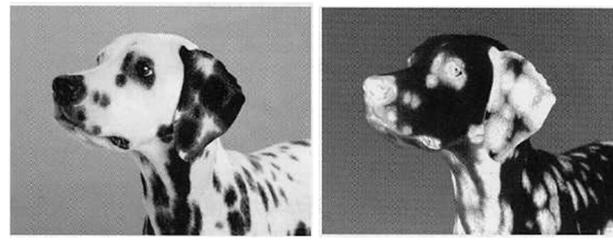**Ritsumeikan University**

Problem 2.   Questions below are about image gray-level transformation and spatial filtering.

(1) Image gray-level transformation is performed on the input images shown in Figure 2.1 and Figure 2.2. The output images are also shown in the same figures. What is the appropriate tone curve used in each figure and what is the transformation type? Choose the correct answer from following options.

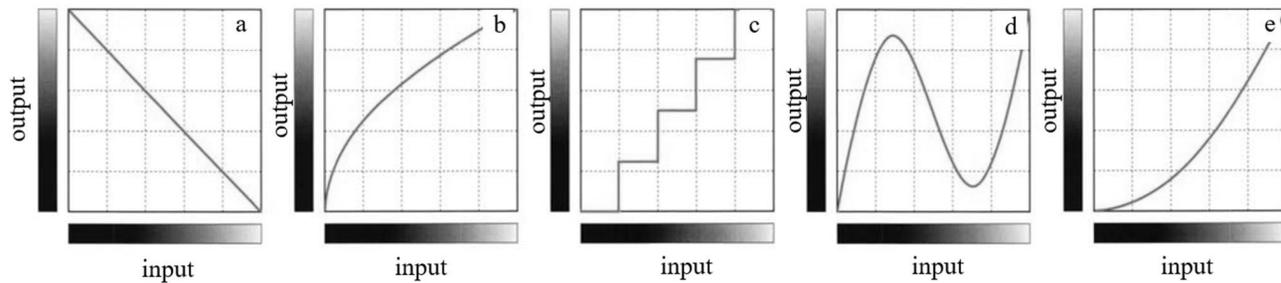| Input | Output | Input | Output |

Figure 2.1   Figure 2.2

【Options】

Tone curves

Transformation types

   a. Gamma correction   b. Black/white inversion   c. Solarization   d. Log transformation   e. Posterization
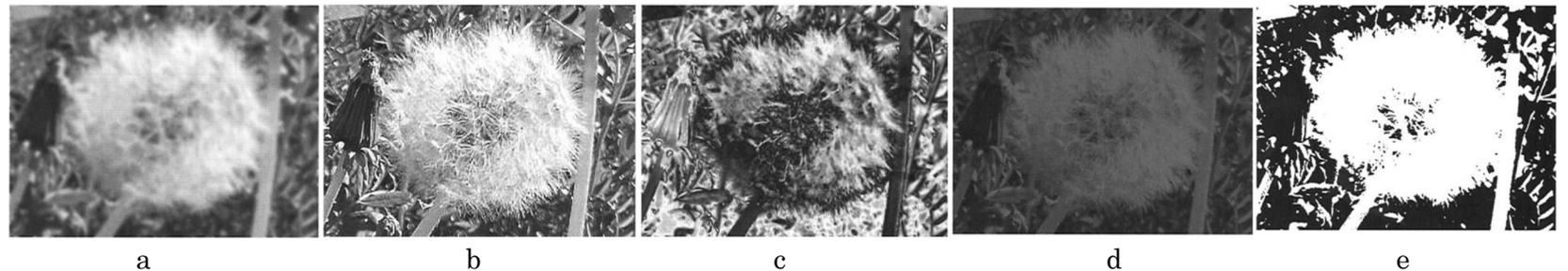
(2) For the grayscale image in Figure 2.3, the following three filters are applied. Choose the correct output image for each filter from the following options.

   ① Sharpening filter
   ② Binary filter
   ③ Smoothing filter

Figure 2.3 original image

【Options】

a   b   c   d   e

Problem 3. Questions below are about geometric transformation

(1) Figure 3(b) is obtained as a result of scaling and skew transformation by applying Equation [1] to Figure 3(a). What are the values of a, c, and d?

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad [1]$$
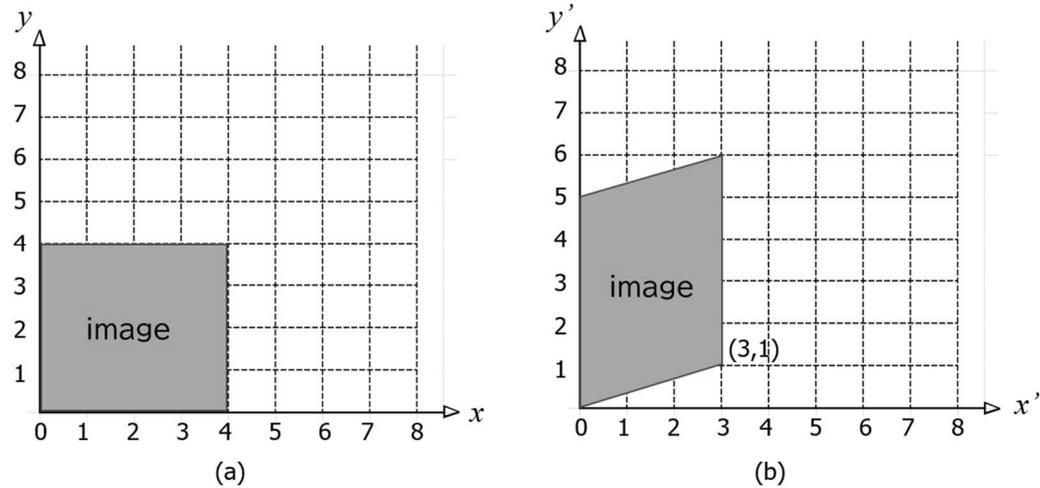

Figure 3

(2) The affine transformation is the combination of arbitrary linear transformation and translation, which is given by Equation [2] in homogeneous coordinates. When rotating (x, y) by 90 degrees counterclockwise and moving one unit on both x and y axes, what are the values of $a \sim f$?

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad [2]$$

(3) When the affine transformation is performed on point $(1, 0)$ under condition (2), what are the values of x' and y' after transformation?

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

Problem 4.   Questions below are about image restoration.

(1) The captured image may be degraded due to "out of focus" or "blur" of the camera. As shown in Equation [3], the degraded image $g(x, y)$ can be considered as the image obtained after applying a degradation filter to the original image $f(x, y)$. In Equation [3], $h(x, y)$ is the degradation filter, and the symbol * represents the convolution integral.

$$g(x, y) = f(x, y) * h(x, y) \qquad [3]$$

If filter $h(x, y)$ is a specific function, the original image $f(x, y)$ will not be degraded and the output is still the original image unchanged. What function is this? Choose the correct answer from following options.

【Options】

a. two-dimensional delta function   b. explicit function   c. point spread function

d. inverse function   e. implicit function

(2) The two-dimensional Fourier transform $G(u, v)$ of the degraded image $g(x, y)$ can be represented as the product of the two-dimensional Fourier transform $F(u, v)$ of the original image, and the two-dimensional Fourier transform $H(u, v)$ of degradation filter $h(x, y)$ (as shown in Equation [4]).

$$G(u, v) = F(u, v)H(u, v) \qquad [4]$$

If the degradation filter $h(x, y)$ is known, which restoration method is correct? Choose the correct answer from following options.

【Options】

a. Multiply $H(u, v)$ by $G(u, v)$ and apply inverse Fourier transform to the result.

b. Multiply $H(u, v)$ by $g(x, y)$ and apply inverse Fourier transform to the result.

c. Multiply the inverse filter of $H(u, v)$ by $G(u, v)$ and apply inverse Fourier transform to the result.

d. Add the inverse filter of $H(u, v)$ to $G(u, v)$ and apply inverse Fourier transform to the result.

e. Multiply the inverse filter of $H(u, v)$ by $g(x, y)$ and apply inverse Fourier transform to the result.

**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.

【1】 **Hamming distance** and **edit distance** are distance measures to evaluate the distance between two strings. Now, we consider measuring the distance between MAMKST and MNAOKST.
（1）Calculate the **Hamming distance** between MAMKST and MNAOKST. Note that the head of the two strings
 i.e., "M" is corresponding and compared to each other.
（2）Calculate the edit distance between "MAMKST" and "MNAOKST" using dynamic programming. Also, please describe the values stored on a matrix via memoization in the dynamic programming procedure.
（3）Describe the computational cost of evaluating the edit distance between two strings whose length are M and N, respectively. Use Big O notation for describing the computational cost.

【2】
（1）Please explain what **(syntactic) parsing** means in natural language processing. Please use the sentence "The boy brings a bottle to the kitchen", as an example for the explanation.

（2）What kind of payoff matrix describes the **prisoner's dilemma** in game theory? Show an example of such a payoff matrix. Also, please explain the meaning of the prisoner's dilemma using the term Nash equilibrium.

（3）Please explain the difference between the **Markov decision process** and the **partially observable Markov decision process**.

（4）Please show the equation of the **Bayesian theorem** and explain the function of the theorem.

（5）Assuming that two sentences "x is a bird" and "x can fly" are represented by **predicate functions** bird(x) and fly(y), respectively, please describe a predicate logical formula representing "all birds can fly."