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</u> or the Human Information Science section.

In case choosing the Computer Science section, answer three questions from question $(4) \sim (9)$. 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	 Linear Algebra Probability and Statistics Data Structure and Algorithms 					
Specialized Subjects	Computer Science	 ④ Computer Architecture ⑤ Operating System ⑥ Software Engineering ⑦ Computer Networks ⑧ Databases ⑨ Artificial Intelligence 				
	Human Information Science	Image Processing Artificial Intelligence				

[Examination time]

9:30-11:30 (120minutes)

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

X 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 Subjects① Linear Algebra

Answer all the questions below.

Question 1. Let $A = \begin{bmatrix} 2 & 0 & 1 \\ -2 & 3 & 4 \\ -5 & 5 & 6 \end{bmatrix}$. (1) Find the inverse of *A*.

(2) Solve the following simultaneous equations using the result of (1).

$$\begin{cases} 4x + 2z = 8\\ -4x + 6y + 8z = 6\\ -10x + 10y + 12z = 4 \end{cases}$$

Question 2. Let $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$.

- (1) Find the eigenvalues and the eigenvectors of A.
- (2) Find the diagonalized result of A.
- (3) Solve A^{50} .

Common Subjects ② Probability and Statistics

Answer all the questions below with derivation.

- Question 1. Find positive constant k of the following joint probability density function f(x, y) for two continuous random variables X and Y. Then, obtain each marginal probability density function for X and Y. Are X and Y independent? Justify your answer.
- (1) $f(x,y) = \begin{cases} k, & (0 < x < 1, & 0 < y < 1), \\ 0, & (0\text{therwise}) \end{cases}$ (2) $f(x,y) = \begin{cases} k \cdot (1-x)(1-y), & (0 < x < 1, & 0 < y < 1), \\ 0, & (0\text{therwise}) \end{cases}$ $(k - (x > 0, & y > 0, & x + y < 1) \end{cases}$
- $(3) f(x,y) = \begin{cases} k, & (x > 0, y > 0, x + y \le 1), \\ 0, & (Otherwise) \end{cases}$
- $(4) \quad f(x,y) = \begin{cases} k \cdot (1-x-y), & (x > 0, y > 0, x+y \le 1), \\ 0, & (\text{Otherwise}) \end{cases}$

(5)
$$f(x,y) = \begin{cases} k \cdot \exp(-x - y), & (x > 0, y > 0), \\ 0, & (Otherwise) \end{cases}$$

Question 2. Suppose X and Y are two discrete random variables that take values from $\{1,2,3\}$, and the joint probability distribution function P(X,Y), (X,Y = 1,2,3) is given by the following table. Answer the questions with derivation.

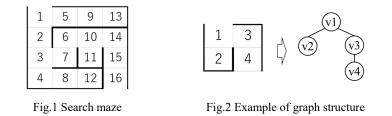
Y X	1	2	3	total
1	1/12	1/6	1/12	1/3
2	1/6	0	1/6	1/3
3	1/12	1/6	1/12	1/3
total	1/3	1/3	1/3	1

- (1) Compute the expectations E[X], E[Y] and E[XY], and show $E[XY] = E[X] \cdot E[Y]$.
- (2) Are X and Y independent? Justify your answer.
- (3) Compute the variances V[X], V[Y], and V[X + Y], and show V[X + Y] = V[X] + V[Y].
- (4) Prove whether the following proposition is correct in general for two random variables X and Y, and give a counterexample if it is not correct: 'If X and Y are independent then $E[XY] = E[X] \cdot E[Y]$ '. How about the converse proposition: 'If $E[XY] = E[X] \cdot E[Y]$ then X and Y are independent'?
- (5) Prove whether the following proposition is correct in general, and give a counterexample if it is not correct: 'If $E[XY] = E[X] \cdot E[Y]$ then V[X + Y] = V[X] + V[Y]'.

Common Subjects③ Data Structure and Algorithms

Answer all the questions below.

Question 1. Answer the following questions when performing a search of the maze shown in Fig.1.



- (1) Refer to Fig.2 and draw Fig.1 as a graph structure. The starting point is assumed to be v1.
- (2) When searching for all v using Breadth-first search, answer the last v to be searched. If there are more than one candidate v, the v with the smallest number is searched first.
- (3) When searching Breadth-first using the Queue given in Fig.3, draw the contents of the Queue just after the dequeue of v16.

• enqueue v1 on an empty Queue
while (Queue is not empty)
• Dequeue from the Queue.
• For the dequeued v, enqueue the neighboring v that have not been searched in the order of ascending number.
Fig.3 Queue based breadth-first search algorithm

- (4) When searching for all v using Depth-first search, answer the last v to be searched. If there are more than one candidate v, the v with the smallest number is searched first.
- (5) When searching Depth-first using the Stack shown in Fig.4, draw the contents of the Stack just after v16 is popped.

• push v1 to the empty Stack
while (Stack is not empty)
• pop from Stack
• For the popped v, push the neighboring v that have not been searched in order of descending number.

Fig.4 Stack based Depth-first search algorithm

Question 2. Quick sort is an algorithm for sorting by repeated division based on elements called pivots. Answer the following questions when sorting by ascending order quick sort on the $S = \{4,7,2,5,8,3,1,9,6\}$. The pivot selection is assumed to be the first element of the number sequence, and the pivot selection is performed when the number of elements in the number sequence is two or more.

- (1) Show the process of a number sequence S being sorted in ascending order by quick sort, using a figure and text.
- (2) There are several orderings of the number sequence S when the total number of pivot choices is maximized. Show any one of them and explain the reason.
- (3) There are several orderings of the number sequence S when the total number of pivot choices is minimized. Show any one of them and explain the reason.

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 $4 \sim 9$

Computer Science ④ Computer Architecture

Answer the question below.

Question:

Survey how current computer memory systems are commonly constructed, using all the technical terms below correctly, and also describing the related technologies. Make a complete explanation <u>using only one side of the answer sheet.</u>

[Technical terms] HDD, RAM, ROM, SSD, access speed, external memory, memory hierarchy, volatile, cache, magnetic, main memory, tape, disk, electric, secondary memory, semiconductor, non-volatile, auxiliary memory, capacity, register

Computer Science ⑤ Operating System

Answer all the questions below.

Question:

There are three cases in which OS codes run. For each case, (a)answer the name of the case, (b)describe the type of timing when the case occurs, and (c)show specific examples of each case. Write the answers in the answer sheet, using the format below.

Case(1)

(a) Name:_____

(b) Timing:

(c)Specific examples:

Case(2)

(a) Name: _

(b) Timing:

(c)Specific examples:

Case(3)

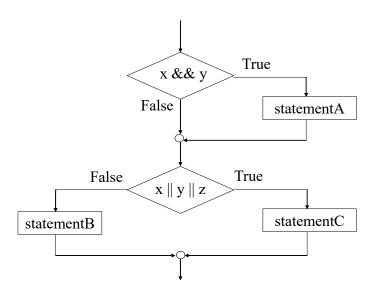
(c)Specific examples:

Computer Science ⑥ Software Engineering

Answer all the questions below.

Question 1.

In the following flowchart, x, y and z are Boolean variables. "&&" is logical AND operation. "||" is logical OR operation. Show the minimum test data that satisfies decision/condition coverage criteria (both branch coverage and codition coverage).



Question 2. Describe two advantages of the function point (FP) methodology in development management.

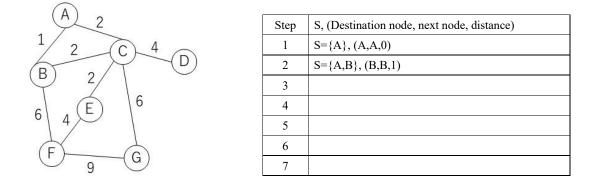
- Question 3. Explain the meaning of "STS decomposition" in structured design.
- Question 4. Explain the meaning of "is-a relation" in an object-oriented approach.
- Question 5. Explain the meaning of "software maintenance."

Computer Science ⑦ **Computer Networks** * This question consists of 2pages.

Answer all the questions below.

Question 1.

Using Dijkstra's algorithm, find the shortest path for routing packets from node A to every other node in the network below. The number near the link denotes the distance for the link. Complete the table below, showing each step of the algorithm. In the second column of the table, write the path set S of nodes to which the algorithm finds the shortest path (destination node, next node, distance). Here, the destination node is the new node in the shortest path set, the next node is the next-hop node from node A to reach the destination node, and the distance is the shortest path distance from node A to the destination node. Write the answers in the answer sheet.



Question 2.

For the descriptions regarding computer networks from (1) to (8), 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) A formal technical specification issued by the IETF, a voluntary organization that promotes the standardization of various

Internet technologies, is Q(a) ① ITU ② RFC ③ ISO ④ IEEE.
(2) Consider the transmission of data from one host A to another host B in the same network. The data transmitted from A to B
is framed by the functions of Q(b) ① physical ② data link ③ network ④ transport layer, and redundant bits are
transmission errors. Q(c) (1) ARP (2) ICMP (3) DNS (4) HTTP is used to obtain A's Q(b) layer address.
(3) Q(d) (1) ASK (2) PSK (3) FSK (4) QAM is a digital modulation process which conveys data by changing the phase
of a constant frequency reference signal (the carrier wave).
(4) A Q(e) (1) bridge (2) repeater (3) router (4) gateway extends a network at the data link layer of the OSI reference
model, and forwards a frame correctly received from one LAN to another LAN.
(5) A DNS record pointing to the IPv4 address of the host is Q(f) (1) A (2) AAAA (3) PTR (4) CNAME record. A DNS
record defining the authoritative name server of a zone is Q(g) ① MX ② NS ③ SOA ④ TXT record.

(6) Regarding TCP and UDP on IPv4, Q(h) ① only TCP retransmits ② only TCP retransmits ③ both TCP and UDP retransmit IP datagrams that are lost due to errors in the IP layer.

(7) The broadcast address for the IPv4 address 192.168.0.0/22 in CIDR is Q(i) ① 192.168.0.1 ②192.168.0.255

3192.168.3.254 4192.168.3.255.

(8) The transport protocol specified in RFC2205 for reserving network resources and real-time communication such as

multimedia information between nodes is Q(j) ① RTP ② RTCP ③ RSVP ④ SCTP.

Computer Science ⑧ Databases

Answer all the questions below.

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

Relations can be normalized based on [①] representing that a value of a set of attributes X determines a value of another attribute A. For example, a relation

Student(<u>student_no</u>, student_name, course_name, course_director_name)

has [1] such that a value of [2] determines a value of *student_name*, and a value of *course_name* determines a value of [3]. In this case, this relation satisfies the condition up to the [4] normal form. In addition, this relation can be decomposed to

R1(*student_no*, *student_name*, [5]) R2([6], [7])

that satisfy the condition of a higher normal form.

[Choices]

(a)	functional	(b)	functional consistency	(c)	functional dependency	(d)	functional atomicity
	uniqueness						
(e)	student_no	(f)	student_name	(g)	course_name	(h)	course_director_name
(j)	first	(k)	second	(m)	third	(n)	Boyce-Codd

Question 2. Three relations are given below.

Student(student_no, student_name)

Subject(subject_no, subject_name)

Score(student_no, subject_no, point)

Choose a correct item from the choices and answer with a symbol (p) to (x) to fill in each of the numbered brackets, for constructing a SQL statement to get an average point of subjects taken by a student whose name is "Taro Ritsumei". (Brackets with the same number are filled by the same item.)

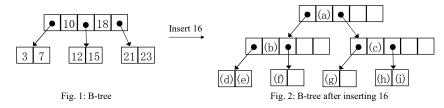
```
SELECT AVG([(\otimes)]) FROM[(\otimes)] WHERE [(w)] IN (SELECT [(w)] FROM[(w)] WHERE [(w)] = "Taro Ritsumei")
```

[Choices]

(p)	Student	(q)	student_no	(r)	student_name	(s)	Subject
(t)	subject_no	(u)	subject_name	(w)	Score	(x)	point

Question 3. B-tree is a data structure which is used for database indices. Answer with a value to fill in each of the numbered brackets in the following sentence.

After value 16 is inserted into B-tree in Fig. 1, (a) is [⁽¹⁾/₃], (c) is [⁽¹⁾/₄] and (g) is [⁽¹⁾/₅] in the B-tree in Fig. 2.



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

Answer all the questions below.

Question : For the following explanations of Artificial Intelligence (AI) from (i) to (v), choose the most suitable word or number for the boxes (1) to (15) from the options and answer with the symbols shown in alphabetic letters. Note that some 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). First-order logic consists of constant symbols, variable symbols, function symbols, predicate symbols, (1), and 5 logical connectives. There are two types of (1): (2) (1) denoted by ∀, and (3) (1) denoted by ∃.
- (ii). A semantic network is a graphical notation of knowledge representation that illustrates concepts and relations between them. A/An (4) relation representing meronymy and holonymy does not have inheritance, while a/an (5) relation representing hyponymy and hypernymy has inheritance.
- (iii). A game where multiple players participate sometimes forms an equilibrium if no player has incentive to switch their strategies. (6) is an equilibrium where each player has a strategy that can increase his/her utility independent of the opponents' actions. On the other hand,

(7) is an equilibrium where each player has a strategy that can increase his/her utility according to the opponents' actions. In the case of two players, utilities acquired by each player in all of the combinations of their actions are arranged in a matrix, named (8). From among the following (1) to (4), (9) results in (6).

	B: action 1	B: action 2		B: action 1	B: action 2		B: action 1	B: action 2		B: action 1	B: action 2
A: action 1	5, 3	3, 5	A: action 1	2, 3	7, 1	A: action 1	4, 4	8, 1	A: action 1	3, 2	8, 1
A: action 2	3, 5	5, 3	A: action 2	5,0	0, 2	A: action 2	1,8	1, 1	A: action 2	5, 4	0, 2
1 2						3			4		

(iv). In a game where two players make decisions in turn, the players can make a strategy using a game tree. (10) is a strategy where the first mover takes an action that increases his/her utility assuming that the second mover tries to decrease the first mover's utility. Pruning is useful to avoid unnecessary search on the game tree in (10). (11) is a pruning of the first mover's action, while (12) is a pruning of the second mover's action. In Figure 1's game tree where the numbers under leaf nodes denote the utilities acquired by the first mover, The (13) edge is pruned as a/an (12).

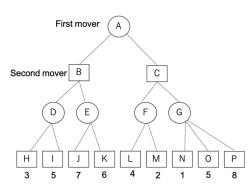


Figure 1: Game tree

(v). Machine learning is classified into three types: supervised learning, unsupervised learning, and

(14). (14) is a learning algorithm that is often modeled based on (15) and can obtain an optimal policy that maximizes the expected future reward.

[Options]

	-								
а	Dominant strategy	b	Quantum symbol	с	Estimated cost	d	Dijkstra's algorithm		
	equilibrium								
е	Null move pruning	f	is-a	g	eta cut	h	neuron		
j	Adjacency matrix	k	Reinforcement	1	union-of	m	has-a		
			learning						
n	Quantifier	0	Markov decision	р	Transfer learning	q	Reward		
			process						
r	Forward pruning	s	Nash equilibrium	t	Universal	u	α cut		
v	Dynamic programming	w	Payoff matrix	х	Perfect	у	Existential		
	algorithm				equilibrium				
Z	do	А	Minimax algorithm	В	Deep learning	D	Operation symbol		
Е	1	G	2	Н	3	J	4		
L	B-D	М	В-Е	Q	C-F	R	C-G		

Specialized Subjects

Human Information Science

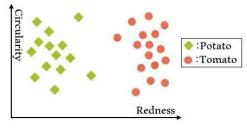
①Image Processing
①Artificial Intelligence

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

* This question consists of 4 pages.

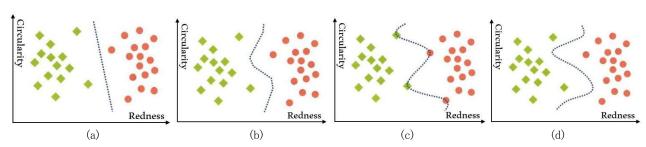
Answer all the questions below.

- Q1. Choose the best answer for questions (1) and (2), and answer with the options from (a) to (e).
- (1) Multiples images of potatoes and tomatoes are used to learn patterns. When redness and circularity are used as features, the feature map can be plotted as the distribution shown in Figure 1. With this learning result, an unknown image can be recognized by the nearest neighbor method using the Euclidean distance. In that case, which of the following is the most suitable decision boundary for discriminating the potatoes and tomatoes?





[Options]



(2) With the learning result shown in Figure 1, a new image is recognized using the *k*-nearest neighbor method (*k* = 3) with Euclidean distance. In Figure 2, three input images A, B, and C are given. Which classes are they recognized as? Choose the best result from the following options.

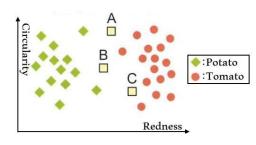


Figure 2

	А	В	С	
(a)	Tomato	Tomato	Tomato	
(b)	Tomato	Tomato	Potato	
(c)	Tomato	Potato	Tomato	
(d)	Potato	Potato	Potato	
(e)	Potato	Potato	Tomato	

[Options]	
-----------	--

- Q2. Choose the best answer for the following (1) and (2) questions, and answer with the option.
- f(x, y) is the original image and g(x, y) is the degraded image. The process of image degradation due to camera blur or out of focus can be expressed by equation ① using the point spread function h(x, y). Here, x, y are the coordinates of images. Which of the following G(u, v) can represent the two-dimensional Fourier transform of g(x, y)? Note that F(u, v), H(u, v) are the two-dimensional Fourier transforms of f(x, y), h(x, y), and * represents the convolution integral.

[Options]

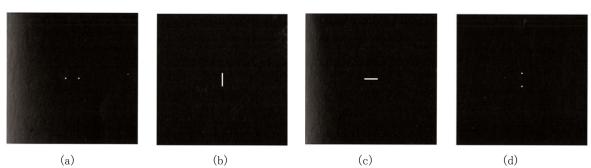
- (a) G(u, v) = F(u, v)H(u, v)
- (b) G(u, v) = F(u, v) * H(u, v)
- (c) $G(u, v) = \frac{H(u, v)}{F(u, v)}$ (d) $G(u, v) = \frac{F(u, v)}{H(u, v)}$
- (2) In equation (1), when f(x, y) is represented as Figure 3(a) and g(x, y) is represented as Figure 3(b), which image can represent h(x, y) in the following options? Here, h(x, y) is normalized so that the sum of its coefficients is 1. The coordinate of the origin is in the center of the image, and 0 is represented by black.



(a)

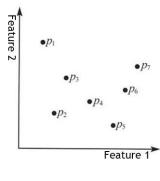


Figure 3

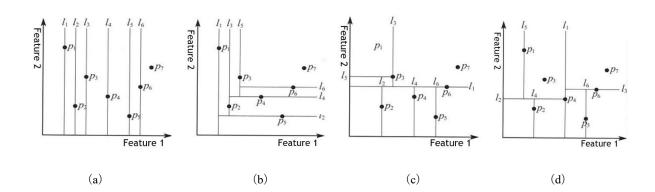


[Options]

- Q3. Choose the best answer for the following (1) and (2) questions, and answer with the option.
- (1) Given 7 two-dimensional data as shown in Figure 4, the purpose is to build a binary tree by the kd-tree method. First, the data are divided by hyperplane l_1 based on feature 1, and divided data are stored in the left and right child nodes. Then, the data stored in the child nodes are sequentially divided by hyperplane l_2 based on feature 2, and continuously divided by the hyperplane based on feature 1, and so on. In the process of building the tree with kd-tree method, which of the following options is the most appropriate hyperplane to separate the data?







(2) Hashing can be used to approximate nearest neighbor. In equation 2, the hash function is represented as h. Which of the following data is the output of the nearest neighbor of test data b = (10110011). Note that, M is 16, and test data b has already been converted to binary vector. The hash values of the data are shown in Table 1.

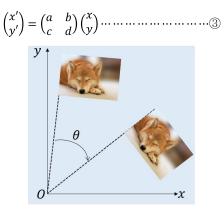
Table 1					
Hash Value	Data				
3	Data 1				
5	Data 2				
9	Data 3				
11	Data 4				
15	Data 5				

[Options]

(Options)

(a) Data 1 (b)Data 2 (c)Data 3 (d)Data 4 (e)Data 5

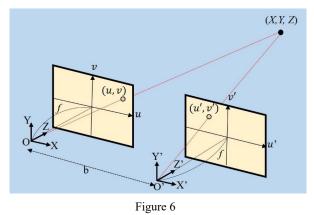
- Q4. Choose the best answer for the following (1) and (2) questions, and answer with the option.
- (1) When rotating an image using the linear transformation with equation (3), which transformation equation can be used to represent rotating the image by θ degrees clockwise about the origin as shown in Figure 5?





[Options]

- (a) $\binom{x'}{y'} = \binom{\sin\theta \cos\theta}{\cos\theta \sin\theta} \binom{x}{y}$ (b) $\binom{x'}{y'} = \binom{1}{\tan\theta} \binom{0}{1} \binom{x}{y}$ (c) $\binom{x'}{y'} = \binom{\sin\theta \cos\theta}{-\cos\theta \sin\theta} \binom{x}{y}$
- (d) $\begin{pmatrix} x'\\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x\\ y \end{pmatrix}$ (e) $\begin{pmatrix} x'\\ y' \end{pmatrix} = \begin{pmatrix} 1 & \tan\theta\\ 0 & 1 \end{pmatrix} \begin{pmatrix} x\\ y \end{pmatrix}$ (f) $\begin{pmatrix} x'\\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x\\ y \end{pmatrix}$
- (2) Which of the following option is the most appropriate to describe the parallel stereo shown in Figure 6, in which two cameras are aligned with an interval b and have a focal length f. Note that the u-axis and u'-axis of the image coordinates lie on the same line and parallel to the X-axis of the world coordinate. And the v-axis and v'-axis of the image coordinates are parallel to the Y-axis of the world coordinate. The origin of (u, v) coordinate lies on the Z axis of the world coordinate, and the parameters of two cameras are the same.



[Options]

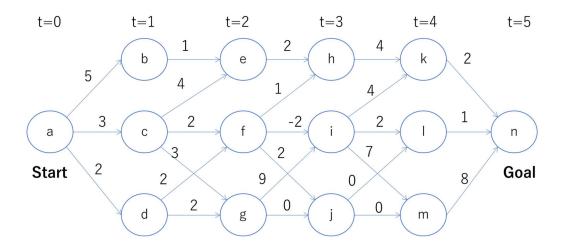
- (a) X is uniquely determined by the parallax u u' and is proportional to the parallax
- (b) X is uniquely determined by the parallax u u' and is inversely proportional to the parallax.
- (c) Y is uniquely determined by the parallax u u' and is proportional to the parallax
- (d) Y is uniquely determined by the parallax u u' and is inversely proportional to the parallax
- (e) Z is uniquely determined by the parallax u u' and is proportional to the parallax
- (f) Z is uniquely determined by the parallax u u' and is inversely proportional to the parallax

Human Information Science 🛈 🛛 Artificial Intelligence

Answer all the questions below.

Question 1. An agent needs to find the path from state \mathbf{a} to state \mathbf{n} that gives the agent the maximum cumulative rewards. The number beside (and just above) each arrow means the reward the agent obtains when the agent moves through the arrow. The cumulative rewards mean the sum of the rewards the agent obtains. Each circle and arrow represent a state and a state transition, respectively. The value t represents a time step. The agent can only move from the left to the right at each step. The problem is to find a path that gives an agent the maximum cumulative reward.

- (1) Execute *dynamic programming* and describe memoized values for all of the states. A memoized value means the maximum cumulative reward to reach the target state (i.e., subgoals).
- (2) Find and describe the optimal path from state **a** to goal **n**.
- (3) When the number of states at each time step becomes M, and the total time steps becomes N, the total computational cost will increase. Please describe the computational cost of *dynamic programming* and explain it.



Question 2.

- (1) Please explain what **Bayes' theorem** means with an example.
- (2) Please explain what a gradient method (or a steepest descent method) in optimization means with an example.
- (3) Please describe the algorithm of k-means clustering.
- (4) Please explain what syntactic parsing in natural language processing means with an example.