

修士課程 社会情報学専攻入学者選抜試験問題
(専門科目)

Entrance Examination for Master's Program
(Specialized Subjects)
Department of Social Informatics

令和 2 年 8 月 1 日 13:00～16:00
August 1, 2020 13:00 - 16:00

【注意】

- ・ 試験開始の合図があるまで中を見てはいけない。
- ・ 問題用紙は表紙を含めて 30 枚である。試験開始後、枚数を確認し、落丁または印刷の不鮮明なものがあれば直ちに申し出ること。
- ・ 問題は 20 題である。このうち第一位の志望区分が指定する条件を満足する 3 題を選択し、解答しなさい。志望区分ごとの指定条件を次ページに示した。
- ・ 解答用紙の表紙に記載されている注意事項についても留意すること。
- ・ 問題 1 題につき、解答用紙 1 枚を使用すること。解答用紙は裏面を使用しても構わないが、使用する場合は裏面に継続することを明記すること。

NOTES

- ・ Do not open the pages before the announcement of the examination's start.
- ・ This is the Question Booklet consisting of 30 pages including this front cover. After the call to start, check that all pages are in order and notify proctors immediately if missing pages or unclear printings are found.
- ・ There are 20 questions. Choose and answer 3 questions in total. The questions you must choose are assigned by your first-choice application group. The list of conditions is given on the next page.
- ・ Read carefully the notes on the front cover of the Answer Sheets, too.
- ・ Use one sheet for each question. If necessary, the reverse side may be used, stating "See verso" at the end of the front page.

第 1 志望区分の問題選択条件

第 1 志望区分	選択条件
社-1a、社-1b、社-2、社-3、 社-5b、社-6、社-14	T1～T5 から 3 題
社-8、社-9	B1～B5 から 3 題
社-10、社-11、社-12	D1～D5 から 3 題
社-13	M1～M5 から 3 題

Questions to be chosen depending on first-choice application group

First-choice Applicant group	Condition of question choice
SI-1a, SI-1b, SI-2, SI-3, SI-5b, SI-6, SI-14	Select three among T1～T5
SI-8, SI-9	Select three among B1～B5
SI-10, SI-11, SI-12	Select three among D1～D5
SI-13	Select three among M1～M5

問題番号 (Number): T-1

1. ある関係スキーマ $R(A_1, A_2, \dots, A_n)$ の上の連言型定義域関係論理問合せは, R 上の定義域関係論理問合せの部分クラスである. 連言型定義域関係論理問合せにおいて R 上の論理式は次により定義される.

- 原子論理式 $R(t_1, t_2, \dots, t_n)$ は論理式である. ここで, t_i ($i = 1, 2, \dots, n$) は変数または定数である.
- φ と ψ が論理式であれば, $\varphi \wedge \psi$ は論理式である.
- φ が論理式, x が φ の自由変数であれば, $\exists x(\varphi)$ は論理式である.

たとえば, 関係スキーマ $S(A, B)$ を考えると,

$$\exists x(S(x, y) \wedge S(x, z))$$

は, S 上の論理式である.

関係スキーマ $R(A_1, A_2, \dots, A_n)$ の上の連言型定義域関係論理問合せは,

$$\{x_1, x_2, \dots, x_m \mid \varphi\}$$

の形をしている. ここで, $m \geq 1$ であり, φ は, $R(A_1, A_2, \dots, A_n)$ の上の論理式である. $\{x_1, x_2, \dots, x_m\}$ は φ の自由変数集合と一致する. また, x_1, x_2, \dots, x_m を出力変数列と呼ぶ.

たとえば, 関係スキーマ $S(A, B)$ を考えると, 次の問合せ q_1, q_2 はいずれも S 上の連言型定義域関係論理問合せである.

$$\begin{aligned} q_1 : & \{w, w \mid \exists v(S(v, w))\} \\ q_2 : & \{y, z \mid \exists x(S(x, y) \wedge S(x, z))\} \end{aligned}$$

変数集合から定数集合への写像を代入と呼ぶ. 関係スキーマ R のインスタンス I , 及び R 上の連言型定義域関係論理問合せ q が与えられたとき, I に対する q の答 $q(I)$ は, 通常 of 定義域関係論理問合せの場合と同様に代入の概念を用いて定義される.

(a) 関係スキーマ $S(A, B)$ のインスタンスとして以下の I を考える. このとき, S 上の連言型定義域関係論理問合せ q_1 の I に対する答 $q_1(I)$ を与えよ. なぜそのような答が得られたかその理由を説明せよ. 説明中では代入の概念を用いること.

$$I$$

A	B
a_1	b_1
a_1	b_2

- (b) 関係スキーマ $S(A, B)$ のインスタンスとして上記 (a) の I を考える. このとき, S 上の連言型定義域関係論理問合せ q_2 の I に対する答 $q_2(I)$ を与えよ. なぜそのような答が得られたかその理由を説明せよ. 説明中では代入の概念を用いること.
- (c) 関係スキーマ $S(A, B)$ 上で関数従属性 $A \rightarrow B$ が成立するならば, 連言型定義域関係論理問合せ q_1 と q_2 は等価になる. その理由を説明せよ.
- (d) 変数集合から変数と定数の集合への写像を拡張代入と呼ぶ. q_2 の変数集合 $\{x, y, z\}$ から q_1 の変数集合 $\{v, w\}$ へ次のような拡張代入 θ が存在する.

$$\theta(x) = v, \quad \theta(y) = w, \quad \theta(z) = w$$

また, この拡張代入により, q_2 の出力変数列は q_1 の出力変数列に写像される. また, θ は, q_2 に現れる原子論理式を q_1 に現れる原子論理式に写像する. 以上のことから, S のどのようなインスタンス I に対しても $q_1(I) \subseteq q_2(I)$ が成立することが結論付けられる. このような結論が得られる理由を説明せよ.

2. 以下の設問に答えよ. ただし, データベースのトランザクション T_i においてデータ x を読み出す操作を $R_i(x)$ で表し, データ x に書き込む操作を $W_i(x)$ で表すものとする.

- (a) データベースのトランザクション集合 $\{T_1, T_2, \dots, T_m\}$ ($m \geq 2$) を考える. $\{T_1, T_2, \dots, T_m\}$ の スケジュール の定義を与えよ.
- (b) 最終状態直列化可能スケジュール の定義を与えよ.
- (c) 次のスケジュールが最終状態直列化可能かどうかを判定せよ. その判定過程を説明せよ.

$$R_1(a)W_2(a)W_2(b)R_3(a)W_1(b)W_3(b)$$

- (d) 上記 (c) のスケジュールの衝突グラフを描き, このスケジュールが衝突直列化可能かどうかを判定せよ. その判定過程を説明せよ.

1. Conjunctive domain relational calculus queries on a relational schema $R(A_1, A_2, \dots, A_n)$ are a subclass of domain relational calculus queries on R . A formula over R for the conjunctive domain relational calculus queries is defined as follows:

- An atomic formula $R(t_1, t_2, \dots, t_n)$ is a formula, where t_i ($i = 1, 2, \dots, n$) is a variable or a constant.
- $\varphi \wedge \psi$ is a formula, where φ and ψ are formulas; and
- $\exists x(\varphi)$ is a formula, where φ is a formula and x is a free variable in φ .

For example, consider a relational schema $S(A, B)$.

$$\exists x(S(x, y) \wedge S(x, z))$$

is a formula over S .

A conjunctive domain relational calculus query over a relational schema $R(A_1, A_2, \dots, A_n)$ has the form

$$\{x_1, x_2, \dots, x_m \mid \varphi\},$$

where $m \geq 1$, and φ is a formula over $R(A_1, A_2, \dots, A_n)$. $\{x_1, x_2, \dots, x_m\}$ is equal to the set of free variables of φ . x_1, x_2, \dots, x_m is called an output variable sequence.

For example, consider a relational schema $S(A, B)$. Both of the following queries q_1 and q_2 are conjunctive domain relational calculus queries over S .

$$\begin{aligned} q_1 : & \{w, w \mid \exists v(S(v, w))\} \\ q_2 : & \{y, z \mid \exists x(S(x, y) \wedge S(x, z))\} \end{aligned}$$

A mapping from a set of variables to a set of variables is called a valuation. Given an instance I on a relational schema R , and a conjunctive domain relational calculus query q over R , the answer $q(I)$ of q on I is defined using the notion of valuation analogously to the case of normal domain relational calculus queries.

- (a) Consider the following instance I on a relational schema $S(A, B)$. Give the answer $q_1(I)$ of the conjunctive domain relational calculus query q_1 over R . Explain the reason why the answer was obtained. Use the notion of valuation in the explanation.

$$I$$

A	B
a_1	b_1
a_1	b_2

- (b) Consider the instance I in the above (a) on the relational schema $S(A, B)$. Give the answer $q_2(I)$ of the conjunctive domain relational calculus query q_2 over R . Explain the reason why the answer was obtained. Use the notion of valuation in the explanation.
- (c) If the functional dependency $A \rightarrow B$ holds on the relational schema $S(A, B)$, the two conjunctive domain relational calculus queries q_1 and q_2 are equivalent. Explain the reason.
- (d) A mapping from a set of variables to a set of variables and constants is called an extended valuation. There exists the following extended valuation θ from the set of variables $\{x, y, z\}$ of q_2 to the set of variables $\{v, w\}$ of q_1 .

$$\theta(x) = v, \quad \theta(y) = w, \quad \theta(z) = w$$

Furthermore, the extended valuation θ maps the output variable sequence of q_2 to the output variable sequence of q_1 . θ also maps an atomic formula in q_2 to an atomic formula in q_1 .

From these observations, we can conclude that $q_1(I) \subseteq q_2(I)$ holds for any instance I on S . Explain the reason why this conclusion was derived.

2. Answer the following questions. Here, we use $R_i(x)$ to denote an operation to read a data x and use $W_i(x)$ to denote an operation to write a data x in a database transaction T_i .
- (a) Consider a set of database transactions $\{T_1, T_2, \dots, T_m\}$ ($m \geq 2$). Give the definition of a schedule of $\{T_1, T_2, \dots, T_m\}$.
- (b) Give the definition of a final state serializable schedule.
- (c) Decide whether the following schedule is final state serializable or not. Explain your decision process.

$$R_1(a)W_2(a)W_2(b)R_3(a)W_1(b)W_3(b)$$

- (d) Draw the conflict graph of the above schedule in (c) and decide whether this schedule is conflict serializable or not. Explain your decision process.

問題番号 (Number): T-2

マルコフ決定過程と強化学習に関する以下の問題(1)-(5)に回答せよ:

(1) 以下の用語を説明せよ:

「ポリシー (方策)」 「マルコフ性」 「効用」 「Q 値」 「モデルフリー」

(2) 「ポリシー反復」とは何か、「価値反復」とはどのように異なるか、具体例を示しつつ、説明せよ

(3) 「Q 学習 (Q-learning)」のアルゴリズムを示せ

(4) 「 ϵ -グリーディ (greedy)法」 「探検関数 (探索関数)」 「後悔 (Regret)」の用語を用いて、強化学習における「探検 (探索、Exploration)」と「利用 (Exploitation)」について説明せよ

(5) 状態空間が大きい問題に対して強化学習が用いられる具体例を挙げ、そのうえで、近似 Q ラーニングを適用するためにどのように状態の一般化を行うと効率的な強化学習が可能になるかを論ぜよ

Answer all the following questions (1)-(5) on Markov decision processes (MDPs) and reinforcement learning:

(1) Explain the following terms:

"policy", "Markov property", "utility", "Q-value", "model-free."

(2) Explain what is a "policy iteration". Explain how it differs from a "value-iteration". Make your own example and use it for the explanation.

(3) Describe the "Q-learning" algorithm.

(4) Explain "exploration" and "exploitation" in reinforcement-learning using terms " ϵ -greedy method", "exploration functions", and "regret."

(5) Make your own example about the use of reinforcement learning for a problem with a large state-space, and then discuss, how you should generalize the states to apply approximate Q-learning in order to enable efficient reinforcement learning.

問題番号 (Number): T-3

整数 m, n が $m \geq n \geq 2$ を満たし、配列 $A[1], \dots, A[n]$ のそれぞれに $\{1, \dots, m\}$ の異なる要素が格納されているものとする。配列 A は $1 \leq i < j \leq n$ なるすべての整数 i, j について $A[i] < A[j]$ が成り立つときかつその時に限りソート済みであると言う。配列がソート済みであるか否かを調べるアルゴリズムをソート検査アルゴリズムと呼び、入力された配列 A がソート済みならば *true* を出力し、そうでないときに *false* を出力するならば、そのソート検査アルゴリズムは正しいという。アルゴリズムの複雑さを以下に定義するランダウの big- O 記法とその類型を用いて評価する。関数 $f, g: \mathbf{N} \rightarrow \mathbf{R}$ に対して、

- $f(n) = O(g(n))$ ($\exists C > 0, \exists n_0$, s.t. $\forall n > n_0, |f(n)| \leq C|g(n)|$ のとき)
- $f(n) = \Omega(g(n))$ ($\exists C > 0, \exists n_0$, s.t. $\forall n > n_0, |f(n)| \geq C|g(n)|$ のとき)
- $f(n) = \Theta(g(n))$ ($\exists C, C' > 0, \exists n_0$, s.t. $\forall n > n_0, C|g(n)| \leq |f(n)| \leq C'|g(n)|$ のとき)

とする。以下の問いに答えよ。

1. 以下のソート検査アルゴリズム Algorithm 1 が正しいことを示せ。また、最悪ケースの複雑さが $\Theta(n)$ であることを説明し、最良ケースの複雑さを示せ。

Algorithm 1 ソート検査

```

1: procedure IsSORTED( $A, n$ )
2:   for  $i = 1$  to  $n - 1$  do
3:     if  $A[i] > A[i + 1]$  then
4:       return false
5:     end if
6:   end for
7:   return true
8: end procedure

```

2. いかなる正しいソート検査アルゴリズムも、その最悪ケースの複雑さは $\Omega(n)$ であることを示せ。アルゴリズムが配列 A の幾つのエントリーにアクセスしなければならないかを考えるとよい。

3. ソート検査アルゴリズム Algorithm 1 の平均ケースの複雑さについて考えるため、 $\{1, \dots, m\}$ の無重複な n 個の要素である $(A[1], \dots, A[n])$ を一様ランダムに選ぶものとする。

- (a) 関数 $f: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ が $\forall i, j \in \{1, \dots, n\}, i \neq j \Rightarrow f(i) \neq f(j)$ のとき単射であるといい、 $\forall i, j \in \{1, \dots, n\}, i < j \Rightarrow f(i) < f(j)$ の

とき単調増加であるという。単射な関数の個数は $m(m-1)\cdots(m-n+1) = \frac{m!}{(m-n)!}$ 個であることを示せ。また、単調増加な関数の個数は $\binom{m}{n} = \frac{m!}{n!(m-n)!}$ 個であることを示せ。

- (b) 前問の結果を用いて A がソート済みである確率を示せ。
- (c) Algorithm 1 において、**for** ループが少なくとも k 回実行される確率を p_k とする。このとき、 $1 \leq k \leq n-1$ において $p_k = 1/k!$ であることを示せ。
- (d) Algorithm 1 における **for** ループの平均実行回数が定数 $e = \exp(1) = \sum_{k=1}^{+\infty} 1/k!$ で抑えられることを示せ。
- (e) 前問の結果を用いて Algorithm 1 の平均ケースの複雑さを求めよ。

4. 以下では n が固定で $n = m$ であるとする。

- (a) ソート検査アルゴリズム Algorithm 1 を改良し、その最悪および平均ケースの複雑さを示せ。ソート済みの配列が $n = m$ の場合にどのようなものになるか考えるとよい。
- (b) セキュリティの観点から、アルゴリズムはあらゆる入力で同一の実行時間、すなわち固定実行時間を持ち、かつ、条件分岐 (**if** 文) を含まないことが望ましい。この条件を満たすソート検査アルゴリズムを $n = m$ の場合について示せ。ただし、整数 x を入力とし、 $x = 0$ のとき *true* を返し、そうでないときに *false* を返す、条件分岐を含まない固定実行時間の関数 *ISZERO* を用いてよい。

In this problem, we are interested in estimating the complexity of *sortedness*, i.e. testing whether an array of integers is sorted. More precisely, we are given an array A of n elements $A[1], \dots, A[n]$ that are *distinct* integers in $\{1, \dots, m\}$, where $m \geq n \geq 2$. We then consider algorithms that test whether the array A satisfies:

$$A[i] < A[j] \quad \text{for all } 1 \leq i < j \leq n.$$

Complexity estimates use Landau's big- O notation and its variants, defined as follows. For $f, g: \mathbf{N} \rightarrow \mathbf{R}$, we write:

- $f(n) = O(g(n))$ when $\exists C > 0, \exists n_0$ s.t. $\forall n > n_0, |f(n)| \leq C|g(n)|$;
- $f(n) = \Omega(g(n))$ when $\exists C > 0, \exists n_0$ s.t. $\forall n > n_0, |f(n)| \geq C|g(n)|$;
- $f(n) = \Theta(g(n))$ when $\exists C, C' > 0, \exists n_0$ s.t. $\forall n > n_0, C|g(n)| \leq |f(n)| \leq C'|g(n)|$.

Answer the following questions.

1. We consider the following algorithm for the sortedness problem.

Algorithm 1 Sortedness test

```

1: procedure ISORTED( $A, n$ )
2:   for  $i = 1$  to  $n - 1$  do
3:     if  $A[i] > A[i + 1]$  then
4:       return false
5:     end if
6:   end for
7:   return true
8: end procedure

```

Show that Algorithm 1 is correct, and has a worst case complexity of $\Theta(n)$. What is its complexity in the best case?

2. Show that any correct algorithm for this problem must have a worst case complexity of at least $\Omega(n)$. Hint: consider the number of entries of the array A that are accessed by the algorithm.
3. We now want to study the complexity of Algorithm 1 *on average*. To do so, we assume that the tuple $(A[1], \dots, A[n])$ is picked uniformly at random among n -tuples of distinct elements of $\{1, \dots, m\}$.
 - (a) A function $f: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ is said to be *injective* when $\forall i, j \in \{1, \dots, n\}, i \neq j \Rightarrow f(i) \neq f(j)$, and to be *increasing* when $\forall i, j \in \{1, \dots, n\}, i < j \Rightarrow f(i) < f(j)$. Show that there are exactly $m(m-1) \cdots (m-n+1) = \frac{m!}{(m-n)!}$ injective functions and $\binom{m}{n} = \frac{m!}{n!(m-n)!}$ increasing functions.

- (b) What is the probability that the array A is sorted? Justify your answer using the previous result.
 - (c) Let p_k be the probability that, in the execution of Algorithm 1, at least k iterations of the **for** loop are carried out. Show that $p_k = 1/k!$ for $1 \leq k \leq n - 1$.
 - (d) Show that the average number of iterations of the **for** loop carried out in Algorithm 1 is bounded above by the constant $e = \exp(1) = \sum_{k=1}^{+\infty} 1/k!$.
 - (e) Deduce the average case complexity of Algorithm 1.
4. We now assume further that n is fixed to $n = m$.
- (a) Propose an improvement of Algorithm 1 in this setting, and give its complexity in the worst case and on average. Hint: you may consider which arrays are sorted when $n = m$.
 - (b) For security reasons, one sometimes wants to use algorithms that have the exact same running time on all inputs, and contain no conditional branches (i.e., no **if** statement). Propose such an algorithm for the sortedness problem when $n = m$. You may assume the existence of a function **ISZERO** that executes in fixed running time without conditional branches, and on input of an integer x , returns *true* if $x = 0$ and *false* otherwise.

問題番号 (Number): T-4

以下の問いに答えよ.

1. ベクトル空間モデル (tf-idf を用いるもの) と Binary Independence モデル (BIM) は, いずれも広く知られた文書検索のためのモデルである. 両者を比較した場合の, それぞれの利点と欠点について考察せよ.
2. 以下で定義される頂点集合 V と辺集合 E からなる有向グラフ G を考える.

$$\begin{aligned} V &= \{a_1, \dots, a_{99}\} \cup \{b_1, \dots, b_{100}\} \\ E &= \{(a_i, a_j) \mid i \neq j, 1 \leq i \leq 99, 1 \leq j \leq 99\} \\ &\cup \{(b_i, b_j) \mid i \neq j, 1 \leq i \leq 100, 1 \leq j \leq 100\} \end{aligned}$$

- (a) G において, 最も高い PageRank スコアを持つ頂点 (複数ある場合はそれら全て) はどれか説明せよ. また, なぜそのような結果になるか考察せよ.
 - (b) G において, 最も高い HITS による authority 度を持つ頂点 (複数ある場合はそれら全て) はどれか説明せよ. また, なぜそのような結果になるか考察せよ.
3. 頂点集合 V と辺集合 E からなる無向グラフ G において, ノード $v \in V$ のクラスタ係数 $C(v)$ は以下のように定義される.

$$C(v) = \frac{v \text{ を含む三角形の数}}{k(v)(k(v) - 1)/2}$$

ただし, $k(v)$ は v の次数である. $k(v) < 2$ の時は, クラスタ係数 $C(v)$ は定義されないものとする. この時, 以下のことを証明せよ.

連結な無向グラフ G の全てのノード v に対して, 「クラスタ係数 $C(v)$ が定義され, $C(v) = 1$ である」が成り立つ時, G は完全グラフである.

Answer the following questions.

1. The vector space model (using tf-idf) and the binary independence model (BIM) are both well-known models for document retrieval. Discuss what are advantages and disadvantages of each method compared with the other.
2. Consider a directed graph G with the following node set V and the edge set E .

$$\begin{aligned} V &= \{a_1, \dots, a_{99}\} \cup \{b_1, \dots, b_{100}\} \\ E &= \{(a_i, a_j) \mid i \neq j, 1 \leq i \leq 99, 1 \leq j \leq 99\} \\ &\cup \{(b_i, b_j) \mid i \neq j, 1 \leq i \leq 100, 1 \leq j \leq 100\} \end{aligned}$$

- (a) Explain which node (if there are ties, all of them) has the highest PageRank score in G . Also discuss its reason.
 - (b) Explain which node (if there are ties, all of them) has the highest HITS authority score in G . Also discuss its reason.
3. Given an undirected graph G with the node set V and the edge set E , the clustering coefficient of $v \in V$, denoted by $C(v)$, is defined as follows:

$$C(v) = \frac{\text{the number of triangles including } v}{k(v)(k(v) - 1)/2}$$

where $k(v)$ is the degree of v . When $k(v) < 2$, $C(v)$ is undefined for v . Prove the following statement.

If every node v in a connected undirected graph G has its clustering coefficient $C(v)$ defined and $C(v) = 1$, G is a complete graph.

問題番号 (Number): T-5

インタフェースの設計と利用について以下の問いに答えよ。

(1) ドナルド・ノーマンが 1988 年に提案した、インタフェースのデザイン原理の 6 項目を全てあげ、各項目を 200 文字以内で述べよ。

(2) インタフェースの評価のためのデータ収集の手法を 5 つあげよ。また、各手法ごとに、それぞれの長所と短所を合わせて 200 文字以内で述べよ。(つまり、(200 文字以内の長所と短所の説明) × 5 項目で合計 1000 文字以内となるように説明せよ。)

(3) GUI(Graphical User Interface)や Touch Interface など、これまでに提案されたインタフェースの種類を 8 個列挙し、それぞれの特徴を 200 文字以内で説明せよ。ただし、GUI や Touch Interface を 8 個の中に含めないこと。

Answer the following questions about user interfaces.

(1) Describe Donald Norman's 6 principles for designing interfaces, proposed in 1988 with less than 100 words each.

(2) Describe 5 methods of data collection for evaluating interfaces, and describe the advantages and disadvantages of each method with less than 100 words (i.e., in total 500 words).

(3) List 8 types of interfaces such as GUI (Graphical User Interface) and Touch Interface, and describe the feature of each type within 100 words. Please note that GUI and Touch Interface shall not be included in the 8 types.

問題番号 (Number): B-1

以下の問題 I および II に答えよ。

- I. ある生物種の形質 X に対する処理 T の効果を検討するために、20 個体について処理前と処理後の形質 X の値を測定した。その結果、処理前： 4.04 ± 2.01 、処理後： 5.21 ± 2.29 、個体ごとの差（処理後－処理前）： 1.11 ± 1.18 となった（数値はいずれも平均値 \pm 標準偏差）。この測定値に対して 2 つの分析（分析 1 および分析 2）を行った。

分析 1：処理前と処理後の平均値に対して等分散を仮定して 2 平均値間の差の検定（t 検定）を行った。その結果有意確率 P は 0.09 となった。

分析 2：処理後－処理前の値の平均値について t 検定を行った。その結果有意確率 P は 0.00045 となった。

- 1) 分析 1 と分析 2 の帰無仮説をそれぞれ示せ。
- 2) 分析 1 と分析 2 の t 分布の自由度をそれぞれ示せ。
- 3) 形質 X に対する処理 T の効果を検討するためには、どちらの分析が適切であるか。

その理由も示せ。また、適切な分析から得られる結論を述べよ。

- II. 50 個体の体重（X1）、内臓脂肪量（X2）および月齢（X3）を測定し、これら 3 形質の相関係数行列を求めた（月齢も連続変数として扱った）。結果を表 1 に示した。さらに月齢（X3）の影響を除いた体重（X1）と内臓脂肪量（X2）の偏相関係数は -0.093 となった。
- これらの結果から、体重（X1）、内臓脂肪量（X2）および月齢（X3）の関係を説明しなさい。

表 1 相関行列

	X1	X2	X3
X1	1.0000000	0.9456760	0.9745993
X2	0.9456760	1.0000000	0.9750623
X3	0.9745993	0.9750623	1.0000000

Answer the following questions.

- I. In order to survey the effects of Treatment T on the trait X, measurements of trait X under pre-treatment and post- treatment were carried out for 20 individuals. Mean and standard deviation for each trial were as follows.

Pre-treatment: 4.04 ± 2.01 , post-treatment: 5.21 ± 2.29 and difference between post- and pre-treatment of each individual: 1.11 ± 1.18 . Two different analyses (Analysis 1 and Analysis 2) were carried out.

Analysis 1 : A t-test under equal variance was used for pre- and post-treatment means. The significance probability was 0.09 for this test results.

Analysis 2 : A paired t-test was used to test the mean difference between pre- and post-treatments of each individual. The significance probability was 0.00045 for this test result.

- 1) Show the null-hypothesis for both Analysis 1 and 2.
- 2) Show the degree of freedom of the t-distribution for Analysis 1 and 2.
- 3) Which analysis is appropriate for the test of the effect of Treatment T on trait X?
Please explain the reason. Describe the conclusion derived from the appropriate analysis.

- II. Body weight (X1), visceral fat (X2) and age in months (X3) of 50 individuals were measured and the correlation coefficients matrix was calculated for these three traits (age in months was also treated as a continuous variable). The result is shown in Table 1. The partial correlation coefficient of X1 and X2, in which the effect of X3 is adjusted, was -0.093. Describe the relationship among body weight (X1), visceral fat (X2) and age in months (X3) using these results.

Table 1 correlation coefficients matrix

	X1	X2	X3
X1	1.0000000	0.9456760	0.9745993
X2	0.9456760	1.0000000	0.9750623
X3	0.9745993	0.9750623	1.0000000

著作権上の理由で非表示

(Likens et al., Ecological Monographs, 40(1), p26, 1970)

上の図は北米の冷温帯地域に位置する落葉広葉樹が優占する森林において、隣接する 2 つの流域の一方の植生を伐採してその影響を観測する実験が行われたときの結果の 1 つを表している。2 つの折れ線は実験期間におけるそれぞれの溪流の流量の変化を示している。↓はどちらか一方の流域で伐採が行われた時点を示している。

1. この実験法の名称を答えなさい。
2. この地域の降水量には顕著な季節変動はなく、平年値で毎月 100 mm 前後の降水が観測される。2 つの流域で観測された流量には季節的な変動が見られるが、この季節変動の原因を説明しなさい。
3. 伐採時点以降、2 つの流域の流量は異なる変化を示している。変化が異なる原因となるメカニズムを伐採が行われた流域を特定して説明しなさい。

Not shown due to copyright

(Likens et al., Ecological Monographs, 40(1), p26, 1970)

An experiment was conducted to cut the vegetation completely in one of two adjacent watersheds and to observe its effects in a forest dominated by deciduous broad-leaved trees in the cool temperate regions of North America. The figure above represents one of the results in that experiment. The two lines show the change in the flow rate of each stream during the experiment. ↓ shows the time when logging was done in one of the watersheds.

1. Give the name of this experimental method.
2. There is no significant seasonal change in precipitation in this area, and monthly average precipitation is around 100 mm. There are seasonal variations in discharge observed in the two watersheds. Explain the cause of this seasonal variation.
3. Since the time of logging, the discharges in the two watersheds show different changes. Explain the mechanisms that cause different changes by identifying the watershed where the logging took place.

問題番号 (Number): B-3

陸上生態系の遷移について、以下の問いに答えなさい。

1. 一次遷移と二次遷移のそれぞれについて記述し、その相違について説明しなさい。
2. 二次遷移の初期段階に系に侵入する植物に多く見られる特徴を 3 つ挙げ、それぞれについてその適応的意義を説明しなさい。
3. 一次遷移に伴う窒素循環の変化を記述しなさい。

Answer the following questions about succession in terrestrial ecosystems.

1. Describe each of primary succession and secondary succession, and explain their differences.
2. Show three characteristics that are commonly found in species in early stage of secondary succession, and explain the adaptive significance for each.
3. Describe the change in nitrogen cycle throughout primary succession.

問題番号 (Number): B-4

動物の行動調査に広く用いられている動物装着型の機器に関する以下の問いに答えなさい。

- 1) 発信機と記録計（データロガー）について、それぞれの長所と短所を論じなさい。
- 2) 電波発信機と超音波発信機について、それぞれの長所と短所を論じなさい。
- 3) 以下の研究を計画するとき、最も適した動物装着型の機器とその適用方法について根拠とともに説明しなさい。
 - a) 繁殖期が終わった後に、アオウミガメがどこに移動し、滞在するかを調べる
 - b) アオダイショウが、1日のうちいつ活動をおこなっているかを調べる

Answer the following questions about animal-borne devices that are widely used in studies of animal behavior.

- 1) Explain advantages and disadvantages of transmitters and data storage tags (data loggers).
- 2) Explain advantages and disadvantages of radio transmitters and ultrasonic transmitters.
- 3) Explain the most appropriate animal-borne devices and their application
 - a) To determine the location where green turtles move to and stay after their reproductive period
 - b) To determine the diel cycle of activity of Japanese ratsnakes

問題番号 (Number): B-5

以下の 6 題の小問から 3 題を選択して解答しなさい。

- (1) 標準偏差と標準誤差の相違について説明しなさい。
- (2) ウィルスと細菌の相違について説明しなさい。
- (3) 生態系サービスを構成するサービスを 2 つ挙げ、その内容を説明しなさい。
- (4) 菌根とは何か、説明しなさい。
- (5) 植物の水利用効率の評価方法を説明しなさい。
- (6) 海洋生物における、プランクトン、ネクトン、ベントスの分類について説明しなさい。

Select and answer three out of the six questions below.

- 1) Explain the difference between standard deviation and standard error.
- 2) Explain the difference between virus and bacteria.
- 3) List two ecosystem services, and explain them.
- 4) Explain what mycorrhizae are.
- 5) Explain methods to evaluate water use efficiency in plants.
- 6) Explain the categorization of plankton, nekton, and benthos, in marine organisms.

問題番号 (Number): D-1

リスク管理と危機管理の目的、プロセス、共通点と相違点について説明せよ。

Explain the purpose, process, commons and differences of risk management and crisis management.

問題番号 (Number): D-2

次の数理計画問題について、以下の問いに答えよ。

Answer the questions on the following mathematical programming problem:

$$\begin{aligned} & \max \log x_1 + a \log x_2 \\ & \text{subject to} \\ & 2x_1 + x_2 \leq 8 \\ & x_1 + x_2 \leq 5 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

- (1) $a=1$ のとき、最適解を示せ。

Solve the problem and show the optimal solution, when $a=1$.

- (2) 最適解が $(x_1, x_2) = (3, 2)$ となる a の範囲を示せ。

Show the interval of a in which the optimal solution of the problem is $(x_1, x_2) = (3, 2)$.

問題番号 (Number): D-3

災害リスク・コミュニケーションとは何か。この概念について具体的な事例を二、三あげて論じなさい。

What is disaster risk communication? Discuss this concept by giving a few concrete examples.

問題番号 (Number): D-4

- (1) GPS 測位における「単独測位」の原理について、以下の語句を用いて説明せよ。
語句：補足すべき衛星の数、航法メッセージ、疑似距離
- (2) 「準天頂衛星システム」について、以下の語句を用いて説明せよ
語句：準天頂軌道、GPS との関係、サブメータ級測位補強サービス、安否確認サービス (Q-ANPI)
- (3) GPS 測位により実現した防災に関連するサービスについて、具体例を 2 つ挙げその内容と効果について説明せよ。

- (1) Explain “stand-alone positioning” in Global Positioning System (GPS) using the following terms.
Terms: Minimum number of satellites for positioning, Navigation Message, Pseudorange
- (2) Explain “Quasi-Zenith Satellite System”, a Japanese satellite positioning system, using the following terms.
Terms: quasi-zenith satellite orbit, Relationship with GPS, Sub-meter Level Augmentation Service, Safety Confirmation Service (Q-ANPI)
- (3) Explain contents and impacts of two concrete GPS based services for disaster prevention and/or disaster response.

問題番号 (Number): D-5

米国の危機管理標準である緊急事態管理システム (NIMS ; National Incident Management System) において、危機対応を成功に導く主要な要素として「Resource Management」がある。「Resource Management」について「情報」の観点から具体例に基づき論じなさい。

The National Emergency Management System (NIMS), a US crisis management standard, is organized into three major components. “Resource Management” is one of the components that lead to successful crisis management. Discuss “Resource Management” from the viewpoints of “information” with a concrete example.

問題番号 (Number): M-1

厚生労働省「医療情報システムの安全管理に関するガイドライン」では、診療録を電子保存する際に、三つの要件を満たすよう求めている。この三つの要件を示し、それぞれどういう意味か述べよ。また、Google Documents などのオンライン文書作成アプリは診療録の作成に利用可能か否かを、前記三要件に照らして述べよ。

“Security Guideline for Medical Information Systems” from Ministry of Medicine, Health, Labour and Welfare requests medical institutions to fulfill three requirements when they store medical records electronically. Explain the three requirements and state their meanings. Also, explain whether online document creation applications such as Google Documents can be used to write medical records under the three requirements.

問題番号 (Number): M-2

特定の治療や検査に対して標準化された患者スケジュールをまとめたものをクリニカルパスといい、医療現場において広く使われている。

- 1) 下記にクリニカルパスの例を示す。「手術前の抗菌薬投与」の有効性を検討するには、どのようにすればよいか、集めるべきデータと、検証の方法について論じよ。

A病院 胆嚢摘出手術クリニカルパス												
	○月○日 手術3日前	○月○日 手術2日前	○月○日 手術前日	○月○日 手術当日	○月○日 手術後	○月○日 術後1日目	○月○日 術後2日目	○月○日 術後3日目	○月○日 術後4日目	○月○日 術後5日目	○月○日 術後6日目	○月○日 術後7日目
食事	普通食	普通食	絶食	絶食	絶食	絶食	お粥	お粥	やわらかい食事		普通食	
清潔	入浴可	入浴可	入浴可	清拭	清拭	清拭	シャワー可	シャワー可	シャワー可	入浴可	入浴可	
活動	制限なし	制限なし	制限なし	絶対安静	絶対安静	絶対安静	リハビリ	リハビリ	リハビリ	リハビリ	リハビリ	退院
検査	採血	採血		採血		採血			採血		採血	
処置				手術								
点滴	抗菌薬	抗菌薬	抗菌薬	抗菌薬		抗菌薬	抗菌薬	抗菌薬	なし			

※抗菌薬：

細菌の生存と繁殖を妨げ、身体に害を与えることを防ぐ薬品

※手術前の抗菌薬投与を正当化する理論：

人間の皮膚には無数の細菌が定着しており、メスで身体を切った際に無菌空間の体内に 菌が侵入するものであるから、事前に抗菌薬を十分に効かせておき、体内での繁殖を妨げることで手術後の感染症（発熱）を防ぐことができる

※手術後感染症：

手術の後に起きる感染症で、細菌が体内で繁殖することで発熱を起こしたり手術痕が膿むなどの症状をきたす

- 2) 1) の分析の結果、術前の抗菌薬投与は無効であると判明した。クリニカルパスをどのように更新すればよいか。また、それにより得られる患者および病院のメリットは何か。医療の質に関するドナベディアン の 3 要素 (Structure, Process, Outcome) に基づいて論じよ。

- 3) クリニカルパスが適用されにくい疾患とはどのようなものか、理由とともに述べよ。

Clinical pathway is a chart healthcare professional regularly uses. The chart contains standardized patient schedule for specific treatment or examination.

1) The chart below is a sample of clinical pathway. When you intend to examine the efficiency of “before-surgery antibiotics”, how do you do it? Discuss what kind of data you have to collect and how to evaluate the efficiency.

Sample Hospital, Clinical Pathway for cholecystectomy												
	Day 1 POD -3	Day 2 POD -2	Day 3 POD -1	Day 4 Before surgery After surgery		Day 5 POD 1	Day 6 POD 2	Day 7 POD 3	Day 8 POD 4	Day 9 POD 5	Day 10 POD 6	Da7 11 POD 7
Meal	Normal meal		Oral intake prohibited				Waterly cooked rice only		Soft meal will be served		Normal meal	
Body Hygiene	Bathing allowed			Bed bath only			Shower allowed			Bathing allowed		
Daily Activity	No limit			Complete rest			Rehabilitation					
Examination	Blood exam	Blood exam		Blood exam		Blood exam			Blood exam		Blood exam	Discharge
Treatment				Surgery								
Intravenous Drip	Antibiotics	Antibiotics	Antibiotics	Antibiotics		Antibiotics	Antibiotics	Antibiotics	None			

* **Antibiotics:** Drug to prevent harmful damage to human body caused by bacteria by killing them or inhibiting them to reproduce.

* **Rationale for before-surgery antibiotics:** Plenty of bacteria colonize on our skin surface. When surgeons make incision, those bacteria will go into patients’ body which must be aseptic (non-bacterial). To prevent surgical-site infection which is caused by bacterial infection, proper administration of antibiotics before surgery is essential.

* **Surgical-site infection:** Bacterial infection that occurs after surgery. Patients develop fever or their wound ulcerate.

2) Analysis on 1) proves before-surgery antibiotics is not effective. How do you update the clinical pathway? Also discuss advantages for patients and hospitals. Refer to Donabedian’s three factors framework of quality of care (Structure, Process, and Outcome) as baseline of your discussion.

3) What kind of diseases does not fit the clinical pathway? Explain the reason.

問題番号 (Number): M-3

肺がんは、癌の中でも最も死亡数が多いがんの一つである。コンピュータ断層撮影(CT)画像を使用し、肺がんを早期発見することができれば、この疾患による死亡を減らすことが期待できる。下記の図では、胸部 CT 画像からの肺結節陰影の検出例が示されている。肺結節陰影は、癌である可能性がある陰影のことで、現状は放射線科医が読影という作業をして見つけ出している。もし、この過程をコンピュータで代替することができれば、肺がんの早期発見につながる。

この過程を機械学習で代替することを考える。一般的な画像識別と異なり、この課題には以下の特徴がある。

- ・一つの胸部 CT 画像の中に複数の肺結節が存在する場合がある
- ・結節は様々な大きさがあり、一定ではない
- ・CT 画像は、RGB 画像と違い、CT 値で構成されている

これらの特徴を踏まえ、課題を解くための適切な機械学習の手法を考え、以下の観点に分けて、その方法を述べよ。

- ・学習データの作り方
- ・識別モデルの設計と学習方法
- ・識別モデルの評価方法

肺結節画像:
著作権上の理由で非表示

Lung cancer is one of the major causes of cancer-related death worldwide. Early diagnosis using computed tomography (CT) images can effectively reduce the mortality. The figure below shows examples of pulmonary nodule detection from lung CT images. Pulmonary nodules are shadows that could be cancer in a lung CT image, and are currently found by radiologists in the process of reading.

In this practice, consider replacing this process with machine learning. Unlike general image classification, this challenge has the following characteristics:

- Multiple lung nodules could be present in one lung CT image
- Nodules vary in size
- CT images are composed of CT values, unlike RGB images

Based on these characteristics, consider an appropriate machine learning method to solve the problem and discuss the method. Also, please provide your answers from the following perspectives.

- Processes to generate training data
- Detection model design and its learning method
- Evaluation method for detection models

Pulmonary nodules image:
Not shown due to copyright

問題番号 (Number): M-4

情報通信技術の発展により、オンラインで医療に関する行為を行う遠隔医療に期待が集まり、平成 30 年 4 月から「オンライン診療料」の診療報酬が新設された。

発展途上の領域であり、様々な診療形式が提案されており、スマートフォンなどのカメラを用いて問診をする想定が多く見受けられる。一方で、カメラだけでなく、超音波エコーや聴診器などの医療機器を介してデータを転送する機能の検討も進んでいる。

情報通信技術の活用によりこのような医療機器の各種データを遠方に届ける際に起こり得る 1) 技術的問題を三つ、2) 運用的問題点を一つ、それぞれ挙げて説明せよ。

Due to the development of information and communication technologies, expectations for telemedicine, which conducts medical-related activities online, have been gathered, and a new health care fee, i.e., "online medical fee", has been established from April 2018. Because it is a developing area, various medical treatment styles have been proposed, and it is often assumed that a medical interview will be conducted using a camera on a smartphone. On the other hand, the function of transferring data via not only the camera but also medical equipment such as ultrasonic echoes and stethoscopes is being developed. Explain 1) three technical issues and 2) one operational issues that may occur when delivering various data generated by this kind of medical devices to a remote site using information communication technology.

問題番号 (Number): M-5

近年、読影レポートの説明漏れが大きな社会問題となっている。読影のプロセスは一般的に、

1. (x 日目) 医師が患者の診察時に放射線検査の必要性を判断し、病院情報システム上で放射線検査の依頼を行う。
2. (y 日目) 患者が放射線検査を受ける。
3. (y 日目) 医師が患者の診察時に病院情報システム上で放射線画像を確認する。
4. (y 日目) 放射線診断医が病院情報システム上で読影レポートを記載する。
5. (z 日目) 医師が患者の診察時に病院情報システム上の読影レポートを患者に説明する。

という流れで行われる。4 の読影レポートの記載は時間がかかり、場合によっては翌日になるため、患者が読影レポートの説明を受けるのは検査を受ける日とは別の日になる。4 の読影レポートで問題が指摘されているにも関わらず 5 の患者への説明が行われなことが読影レポートの説明漏れという問題である。この問題の原因はどこにあるか、情報システムはこの問題に対しどのような対応ができるか、また、情報システムによる対応の限界は何か考察せよ。

Recently, failing to explain contents of radiology reports has become a big social problem. The process of radiology reporting is generally as follows:

1. (day x) A doctor evaluates the necessity of a radiology test and requests the radiology test when seeing a patient.
2. (day y) The patient has the radiology test.
3. (day y) The doctor checks the images of the radiology test on a hospital information system when seeing the patient.
4. (day y) A radiologist writes a radiology report on the hospital information system.
5. (day z) The doctor explains the contents of the radiology report when seeing the patient.

Because the writing of a radiology report in step 4 takes a lot of time and it may be written on the day after day y, the day when a doctor explains the contents of a radiology report and the day when a patient undertakes a radiology test are usually different. Sometimes a problem happens where, even though the radiology report points out the issue with the patient, a doctor fails to explain it to the patient properly. Discuss where the cause of this problem is, what measures information systems could take to address it. At the same time, explain what are limitations of utilizing the information systems for addressing the problem.