

修士課程 社会情報学専攻入学者選抜試験問題 (情報学基礎)

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

令和元年 8 月 5 日 10:00~12:00
August 5, 2019 10:00 - 12:00

【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 6 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 5 題である. このうち 3 題を選択し, 解答しなさい.
- ・ 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 of 6 pages including this front cover.
After the call of starting, check all pages are in order and notify proctors immediately if missing pages or unclear printings are found.
- ・ There are 5 questions. Choose and answer 3 questions in total.
- ・ 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 page.

問題番号（Number）1

データストレージに関する以下の問い合わせに答えよ。

Answer the following questions regarding data storage.

- (1) 磁気ディスクによるディスクストレージシステムの構造について以下の語を全て用いて説明せよ。

Explain the structure of disk storage systems on magnetic disks using all of the following terms.

読み書きヘッド, ランク, シリンダ
read/write head, track, cylinder

- (2) フラッシュドライブが磁気ディスクよりも優れている点を1つ述べよ。

Describe one advantage that flash drives have over magnetic disks.

- (3) データを格納する際にはデータ圧縮が有用なことが多く、データ圧縮手法にはロスなしの手法とロスありの手法がある。ロスなしデータ圧縮技法を1つ挙げて説明せよ。

Data compression is often helpful for storing data, and there are lossless schemes and lossy schemes for data compression. Pick up one lossless data compression technique and explain it.

問題番号 (Number) 2

プロセス間の競合に関する以下の問い合わせに答えよ.

Answer the following questions regarding competitions among processes.

(1) セマフォとは何か説明せよ.

Explain what a semaphore is.

(2) デッドロックの発生条件を説明せよ.

Explain the conditions under which a deadlock occurs.

(3) デッドロックの回避手法を 1 つ説明せよ.

Explain one technique for avoiding deadlocks.

問題番号 (Number) 3

ネットワークとインターネットに関する以下の問い合わせに答えよ。

Answer the following questions on networking and the Internet.

- (1) ネットワークを分類する基準であるネットワークトポロジーにはいくつかの種類がある。よく利用されるトポロジーであるバス型とスター型について述べ、それらの違いを説明せよ。

There are several types of network topologies that are the basis for classifying networks. Describe the two popular ones: the bus topology and the star topology and explain their difference.

- (2) ウェブ上のドキュメントを探索してどこに存在するかを発見するために、各ドキュメントにはユニフォームリソースロケータ(Uniform Resource Locator, URL)という一意なアドレスが与えられている。URLの主要な要素を特定し、以下のURLを例に用いてそれらの意味を簡潔に説明せよ。

In order to locate and retrieve documents on the Web, each document is given a unique address called a Uniform Resource Locator (URL). Please identify the key segments and briefly explain their meaning in the following example of URL:

<https://bankofisland.com/data/files/clientresource.html>

- (3) インターネットソフトウェア階層構造の4つの層を列挙し、各層によって実行される動作をそれぞれ説明せよ。

List the four layers in the Internet software hierarchy and identify a task performed by each layer.

問題番号 (Number) 4

アルゴリズムに関する以下の問い合わせに答えよ。

- (1) 挿入法による整列アルゴリズムを擬似コードで記述せよ。
- (2) 再帰構造を使用するアルゴリズムの例を 1 つ挙げ、擬似コードを用いて記述せよ。

Answer the following questions on algorithms.

- (1) Write a pseudocode for the insertion sort algorithm.
- (2) Give an example of an algorithm that uses recursive structures, and write a pseudocode for the algorithm.

問題番号 (Number) 5

データ構造に関する以下の問い合わせよ.

Answer the following questions on data structures.

(1) リスト, スタック, キューの違いを要約せよ.

Summarize the distinction between lists, stacks and queues.

(2) コンピュータのメインメモリ中にリストを実装する方法として, 連續リストと連結リストという 2 つの基本的な格納構造がある. 以下の問い合わせよ.

There are two basic storage structures for implementing lists in computer memory:
contiguous list and linked list. Answer the following:

a. 連續リストと連結リストで要素の削除がどのように行われるか, それぞれ例を図示して説明せよ.

Draw an example diagram of the contiguous list and the linked list and use it to explain how deleting an entry is performed in these two types of structures.

b. 連續リストは静的なリストを実装するのにより適切であり, 動的なリストを実装するには不適切だと思われるるのはなぜか?

Why is a contiguous list considered more convenient to implement static lists but not for implementing dynamic lists?

(3) データ型と型のインスタンスの違いは何か?

What is the difference between a data type and an instance of that type?

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

令和元年 8 月 5 日 13:00～16:00
August 5, 2019 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 23 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 23 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～T6 から 3 題 |
| 社-8、社-9 | B1～B4 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～T6 |
| SI-8, SI-9 | Select three among B1～B4 |
| 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

- (a) データベースにおけるデータ独立性の概念について説明せよ.
- (b) ある大学に関するデータベーススキーマを設計したい. 次の設問に答えよ.
- (i) 次の四つの条件を満足するデータベーススキーマ S の実体関連図 (ER 図) を描きなさい. ER 図で使用する記法は説明すること. それぞれの実体型と関連型が大学のどのような情報を表すかを説明すること.
- S は三つ以上の実体型を持つ.
 - S の各実体型は三つ以上の属性を持つ.
 - S は三つ以上の関連型を持つ.
 - S の少なくとも一つの関連型は属性を持つ.
- (ii) (i) の実体関連スキーマを関係データベーススキーマに変換せよ. 変換の過程及び変換後の関係データベーススキーマを詳しく記述せよ. 変換後の各関係スキーマの正規形についても説明すること.
- (c) 関係スキーマ $R(A)$ を考える. 属性 A の定義域は整数とする. このとき, 属性 A に現れる整数のうち 2 番目に大きいものを求める問合せを (i) 関係代数と (ii) 関係論理のそれぞれを用いて表現せよ.
- (d) データベースシステムにおける索引の重要性を例を用いて説明せよ.
- (e) ロックを用いたトランザクションの並行制御法について説明し, その方法の利点と欠点についても説明せよ.

- (a) Explain the notion of data independence in databases.
- (b) We want to design a database schema of a university. Answer the following questions.
- (i) Draw an entity-relationship diagram (ER diagram) of a database schema S which satisfies the following four conditions. Explain the notations used in the ER diagram. Explain what information of the university each entity type and relationship type represent.
- S has three or more entity types;
 - each entity type in S has three or more attributes;
 - S has three or more relationship types; and
 - at least one of the relationship types in S has an attribute.
- (ii) Translate the entity-relationship schema in (i) into a relational database schema. Give a detailed description of the translation process and the relational database schema after translation. Explain also the normal form of each relational schema after the translation.
- (c) Consider a relational schema $R(A)$ where the domain of the attribute A is the set of integers. Write a query to answer the second largest integer which appears in the attribute A in (i) relational algebra and (ii) relational calculus.
- (d) Explain the importance of indexes in database systems using examples.
- (e) Explain the method for concurrency control of transactions using locks. Explain also merits and demerits of the method.

問題番号 (Number): T-2

確率的推論に関する以下の問題(1)-(4)に回答せよ:

- (1) 以下の用語を説明せよ:
「不確かさ」「条件付き独立性」「列挙による推論」「ベイズの法則」「マルコフ性」
- (2) 「隠れマルコフモデル」とは何か、「マルコフモデル」とはどのように異なるか、具体例を示しつつ、説明せよ
- (3) 「パーティクルフィルター」のアルゴリズムを示せ
- (4) 人工知能に関する諸課題（例えば、コンピュータチェス、地図の塗分け、音声認識、自己位置同定、など）において確率的推論の手法を選択すべきなのは一般的にどのような場合か、幾つかの課題例を挙げながら議論せよ。その際、「どういう条件が成立すると推論が可能になるか」を明確にすること。

Answer all the following questions (1)-(4) on probabilistic reasoning:

- (1) Explain the following terms:
"uncertainty", "conditional independence", "inference by enumeration", "Bayes rule", "Markov property".
- (2) Explain what a "Hidden Markov Model" is. Explain how it differs from a "Markov Model".
Make your own example and use it for the explanation.
- (3) Describe the "particle filtering" algorithm.
- (4) Among various problems in Artificial Intelligence (for example, computer chess, map coloring, speech recognition, self-localization, and so on), in general, when should we select probabilistic reasoning techniques? Discuss it with reference to a few example problems. Make clear "what condition(s) need(s) to be satisfied in order for reasoning to be possible."

問題番号 (Number): T-3

パーセプトロンとサポートベクターマシンに関する以下の 6 つの問い合わせに答えよ。

- (1) 表 1 に示す 4 つのデータ点を持つデータセットに対して単純パーセプトロンの学習を行う。表 1 に示す分類問題において、 x_1, x_2 は入力を表し、 y は望ましい出力を表す。ここで x_1, x_2 は実数変数であり、 y はブール変数である。パーセプトロンの学習において、データ点は d1, d2, d3, d4 という定められた順番で調べられる。また、 w_1, w_2 はそれぞれ入力 x_1, x_2 に対する重みを表し、 w_0 はバイアス項を表す。重み w_1, w_2 とバイアス w_0 は、最初はすべて 0 に設定されている。重み w_1, w_2 とバイアス w_0 がどのように更新されるか示せ。また、最後に得られる分類器の w_1, w_2, w_0 の値を答えよ。

表 1

| | x_1 | x_2 | y |
|----|-------|-------|-------|
| d1 | 1 | 4 | True |
| d2 | 2 | 3 | True |
| d3 | 4 | 5 | False |
| d4 | 5 | 6 | False |

- (2) サポートベクターマシン(SVM)を考える。表 1 に示すデータセットに対して SVM の学習を行うとき、最後に得られる分類器の w_1, w_2, w_0 の値を答えよ。値を求める過程についても説明すること。
- (3) 学習過程の比較、および、問(1), (2)で得られた分類器の比較により、単純パーセプトロンと SVM の優劣を説明せよ。
- (4) カーネル関数を用いた SVM は線形分離可能でないデータを持つ分類問題を扱うことが可能となる。カーネル関数について説明せよ。
- (5) パーセプトロン学習にカーネル関数を導入することを考える。これはカーネルパーセプトロンと呼ばれる。カーネルパーセプトロンでは、内積がカーネル関数に置き換えられる。まず、単純パーセプトロンにおいて n 個のデータ点を調べた後で、重みベクトルがどのように表されるか答えよ。つぎに、カーネルパーセプトロンの定式化を答えよ。
- (6) 問(5)の解を観察することで、SVM と比較してカーネルパーセプトロンの欠点を説明せよ。

Answer the following six questions about perceptron and support vector machines.

(1) You are training a perceptron on a dataset with four data points shown in Table 1. Table 1 shows a classification problem with x_1 , x_2 representing the inputs and y representing the desired output. Here, x_1 , x_2 are real number variables, and y is a Boolean variable. In learning the perceptron, the data points are examined in the given order, d1, d2, d3, and d4. Also, w_1 , w_2 represent the weights corresponding to x_1 , x_2 respectively, and w_0 represents the bias term. Assume that all the weights w_1 , w_2 and bias w_0 are set to zero at the beginning. Show how the weights w_1 , w_2 and bias w_0 are updated, and write the values of w_1 , w_2 , w_0 of the finally obtained classifier.

Table 1

| | x_1 | x_2 | y |
|----|-------|-------|-------|
| d1 | 1 | 4 | True |
| d2 | 2 | 3 | True |
| d3 | 4 | 5 | False |
| d4 | 5 | 6 | False |

(2) Consider a support vector machine (SVM). Write the values of w_1 , w_2 , w_0 of the finally obtained classifier, when you are training the SVM on a dataset shown in Table 1. Also, explain the process of finding the values.

(3) Explain the pros and cons of a perceptron and SVMs by comparing the classifiers obtained in questions (1) and (2) as well as their learning processes.

(4) SVMs with kernel functions can deal with classification problems having non-linearly separable data. Explain kernel functions.

(5) Consider introducing a kernel function to a perceptron learning, which is called kernel perceptron. In kernel perceptron, the inner product is replaced by a kernel function. First, answer how the weight vector can be represented after examining n data points in a perceptron. Second, write the formulation of kernel perceptron.

(6) Explain the disadvantages of kernel perceptron compared to SVMs by observing the answer to question (5).

問題番号 (Number): T-4

要素N個の要素からなる集合 S 上のランダムオラクル H を、非負整数を入力として以下の条件に従って S の要素を出力するアルゴリズムとする。

- 入力された非負整数 x に対して過去に一度も H が呼び出されていない場合、H は S から一様独立に要素を選び、その値を H(x) として出力する。
- 入力された非負整数 x に対して過去に H が呼び出されている場合、S の要素 H(x) を出力する。

$x \neq y$ および $H(x) = H(y)$ を満たす非負整数の対 (x, y) を H の衝突と呼ぶ。以下の問い合わせよ。

1. ランダムオラクル H が重複のない q 個の非負整数 x_1, \dots, x_q を入力として与えられ、値 $H(x_1), \dots, H(x_q)$ を出力したとする。この時、出力値が重複を含まない確率は

$$P_q = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{q-1}{N}\right)$$

であることを示せ。

2. 非負の値 z に対して不等式 $1 - z \leq \exp(-1)$ が成り立つことを利用して、

$$P_q \leq \exp\left(-\frac{q(q-1)}{2N}\right)$$

であることを示せ。また、 q_0 を $q_0 \geq 1 + \sqrt{4N \log N}$ を満たす最小の整数とするとき、 $q \geq q_0$ である全ての q に対して $P_q \leq \frac{1}{N^2}$ となることを示せ。

3. H の衝突を探索する以下のアルゴリズム（アルゴリズム 1）を考える。このアルゴリズムは S の要素を値とし、0 から N までの整数をインデックスとする配列 A に読み書きができる。インデックス i に格納された値を $A[i]$ と書くことにする。

アルゴリズム 1 が常に停止し、かつ、1 つの衝突 (x, y) を出力することを説明せよ。

Algorithm 1 Collision search

```

1: Procedure FINDCOLLISION()
2:   A[0]←H(0)
3:   for y=1 to N do
4:     s←H(y)
5:     for x=0 to y-1 do
6:       if A[x]=s then return (x,y)
7:     end for
8:     A[y]←s
9:   end for
10: end procedure

```

4. アルゴリズム 1 がランダムオラクル H を $q+1$ 回呼び出して停止する確率は $P_q \cdot q/N$ であることを示せ。また、停止するまでの H の呼び出し回数の期待値が

$$\sum_{q=1}^N P_q \cdot \frac{q(q+1)}{N}$$

であることを示せ。

5. 問 2 及び 4 から、停止するまでの H の呼び出し回数の期待値は高々 $O(\sqrt{N} \log^{3/2} N)$ であることを示せ。
6. アルゴリズム 1 が H を q 回呼び出して停止した時、その実行時間は $\Theta(q^2)$ であることを説明せよ。適切なデータ構造を用いてアルゴリズム 1 を変更することで、実行時間を $O(q \log q)$ に低減できることを説明せよ。

In this problem, we consider a few properties of *random oracles*. We fix a set S with N elements, and define a random oracle H to be a procedure that receives non-negative integers as queries, and replies with elements of S as follows:

- on input of a non-negative integer x never queried before, the random oracle returns an element $H(x)$ of S , picked uniformly at random and independently of all previous replies;
- on input of a non-negative integer x queried previously, it returns the same element $H(x)$ of S as before.

We propose to construct and analyze algorithms to find *collisions* on a random oracle H . A collision on H is any pair (x, y) of non-negative integers such that $x \neq y$ and $H(x) = H(y)$.

Reply to the following questions.

1. Suppose the random oracle is queried on distinct non-negative integers x_1, \dots, x_q , and returns the values $H(x_1), \dots, H(x_q)$. Show that the values $H(x_1), \dots, H(x_q)$ are all distinct with probability exactly:

$$P_q = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{q-1}{N}\right).$$

2. Using the inequality $1 - z \leq \exp(-z)$ ($z \geq 0$), prove that:

$$P_q \leq \exp\left(-\frac{q(q-1)}{2N}\right).$$

Let q_0 be the smallest integer such that $q_0 \geq 1 + \sqrt{4N \log N}$. Show that for $q \geq q_0$, we have $P_q \leq 1/N^2$.

3. Consider the following algorithm, which tries to find collisions on the random oracle H . It is assumed that the algorithm can read and write to an array A of elements of S with indices between 0 and N . The array element with index i is denoted by $A[i]$.

Algorithm 1 Collision search

```

1: procedure FINDCOLLISION()
2:    $A[0] \leftarrow H(0)$ 
3:   for  $y = 1$  to  $N$  do
4:      $s \leftarrow H(y)$ 
5:     for  $x = 0$  to  $y - 1$  do
6:       if  $A[x] = s$  then return  $(x, y)$ 
7:     end for
8:      $A[y] \leftarrow s$ 
9:   end for
10: end procedure
```

Show that Algorithm 1 always terminates, and always outputs a correct collision (x, y) .

4. Show that Algorithm 1 terminates after exactly $q + 1$ oracle queries with probability $P_q \cdot q/N$. Deduce that the expected number of oracle queries is given by:

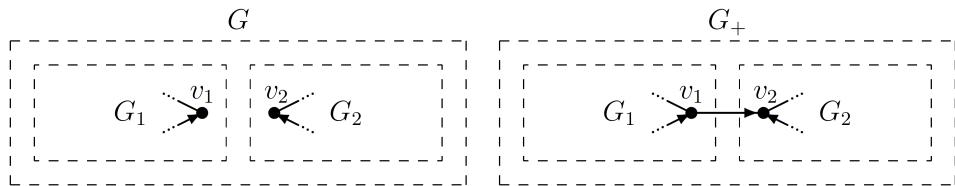
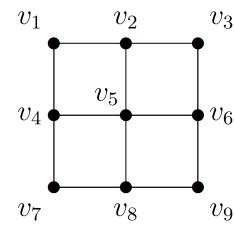
$$\sum_{q=1}^N P_q \cdot \frac{q(q+1)}{N}.$$

5. Using the results of Questions 2 and 4, show that the expected number of oracle queries is at most $O(\sqrt{N} \log^{3/2} N)$.
6. Show that if Algorithm 1 terminates after q queries, its running time is $\Theta(q^2)$. Discuss how one could modify Algorithm 1 using a suitable data structure in order to reduce the running time down to $O(q \log q)$.

問題番号 (Number): T-5

以下の問い合わせよ.

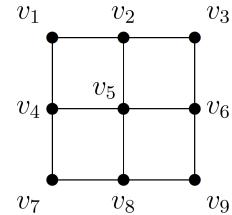
1. 情報検索で用いられる F_1 スコア (balanced F-measure) の定義を示せ. また, F_1 スコアを定義する際に, 適合率と再現率の単純な算術平均を用いない理由を説明せよ.
2. 右下図のような v_1, \dots, v_9 の 9 個のノードからなる無向グラフを考える. このグラフに一つだけ辺を追加するとする.
 - (a) どの頂点間に辺を追加すれば, 左上のノード v_1 の近接中心性が最も大きく上昇するか. 同点となる辺がある場合は, それら全てを示せ.
 - (b) どの頂点間に辺を追加すれば, 左上のノード v_1 の媒介中心性が最も大きく上昇するか. 同点となる辺がある場合は, それら全てを示せ.
3. 頂点集合が V_1 , 辺集合が $E_1 \subseteq V_1 \times V_1$ である有向グラフ $G_1 = (V_1, E_1)$ と, 頂点集合が V_2 , 辺集合が $E_2 \subseteq V_2 \times V_2$ である有向グラフ $G_2 = (V_2, E_2)$ とが与えられているとする. ただし, V_1 と V_2 は互いに素, すなわち, $V_1 \cap V_2 = \emptyset$ であるとする. また, 下図左のような G_1 と G_2 の和にあたるグラフ $G = (V_1 \cup V_2, E_1 \cup E_2)$ を考える. さらに, $v_1 \in V_1$ と $v_2 \in V_2$ を選び, グラフ $G_+ = (V_1 \cup V_2, E_1 \cup E_2 \cup \{(v_1, v_2)\})$ を考える. ここで, 辺 (v_1, v_2) は v_1 から v_2 への有向辺である. すなわち, G_+ は下図右のように G に有向辺 (v_1, v_2) を追加したグラフである.



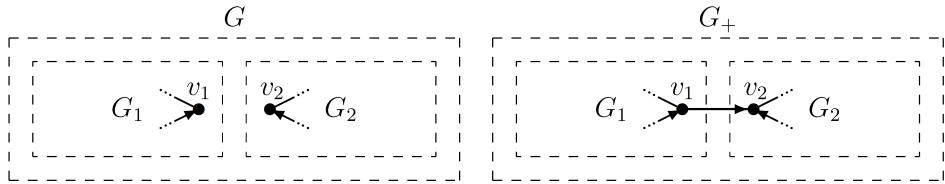
- (a) G と G_+ のそれぞれにおいて, PageRank アルゴリズムを用いて, 各頂点の PageRank スコアを求めた場合, G における v_1 の PageRank スコアと G_+ における v_1 の PageRank スコアとの間にはどのような関係があるか述べよ. また, その理由を説明せよ. ただし, damping factor (teleportation probability) d は, $0 < d < 1$ を満たすものとする.
- (b) G と G_+ のそれぞれにおいて, HITS アルゴリズムを用いて, 各頂点の authority スコアを求めた場合, G における v_1 の authority スコアと G_+ における v_1 の authority スコアとの間にはどのような関係があるか述べよ. また, その理由を説明せよ. ただし, HITS アルゴリズムにおいて authority スコアの正規化は行わない, すなわち, 全ての authority スコアをそれらの最大値で割ることはしないものとする.

Answer the following questions.

1. Give the definition of the F_1 score (balanced F-measure) used in information retrieval. Also explain why a simple arithmetic mean of the precision and recall is not used in the definition of the F_1 score.
2. Consider an undirected graph shown below to the right, which consists of 9 nodes v_1, \dots, v_9 . Suppose we add one edge to this graph.
 - (a) Which edge leads to the largest increase of the closeness centrality of v_1 ? If there are ties, show all edges that tie.
 - (b) Which edge leads to the largest increase of the betweenness centrality of v_1 ? If there are ties, show all edges that tie.



3. Suppose we have a directed graph $G_1 = (V_1, E_1)$, where V_1 is its node set and $E_1 \subseteq V_1 \times V_1$ is its edge set, and a directed graph $G_2 = (V_2, E_2)$, where V_2 is its node set and $E_2 \subseteq V_2 \times V_2$ is its edge set. V_1 and V_2 are disjoint, i.e., $V_1 \cap V_2 = \emptyset$. We also consider $G = (V_1 \cup V_2, E_1 \cup E_2)$ shown below to the left, which corresponds to the union of G_1 and G_2 . In addition, we choose $v_1 \in V_1$ and $v_2 \in V_2$, and consider $G_+ = (V_1 \cup V_2, E_1 \cup E_2 \cup \{(v_1, v_2)\})$ shown below to the right, where (v_1, v_2) is a directed edge from v_1 to v_2 . In other words, G_+ is a graph produced by adding an directed edge (v_1, v_2) to G .



- (a) Suppose we run the PageRank algorithm on each of G and on G_+ , and compute the PageRank scores of their nodes. What relationship exists between the PageRank score of v_1 in G and the PageRank score of v_1 in G_+ ? Also explain why that relationship exists. Assume that we use a damping factor (teleportation probability) d such that $0 < d < 1$.
- (b) Suppose we run the HITS algorithm on each of G and on G_+ , and compute the authority scores of their nodes. What relationship exists between the authority score of v_1 in G and the authority score of v_1 in G_+ ? Also explain why that relationship exists. Assume that we do not normalize the authority scores of nodes in the HITS algorithm. That is, we do not divide the authority scores by the largest one.

問題番号 (Number): T-6

ユーザインターフェースの設計と利用について以下の問い合わせよ。

- (1) ドナルド・ノーマン(Donald Norman) が 1988 年に提案した、ユーザインターフェースのデザイン原理(Norman's design principles)の 6 項目を全てあげ、各項目を 200 文字以内で述べよ。
- (2) GUI(Graphical User Interface)や Voice Interface など、ユーザインターフェースの種類を 5 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 Voice Interface を 5 個の中に含めないこと。
- (3) ユーザインター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。)

Answer the following questions about the user interface.

- (1) Describe Donald Norman's 6 principles for designing user interfaces, proposed in 1988, with less than 100 words each.
- (2) List 5 types of user interfaces such as GUI (Graphical User Interface) and Voice Interface, and describe the feature of each type within 100 words. Note that GUI and Voice Interface should not be included in the 5 types.
- (3) List 5 methods of data collection for evaluating the user interfaces and describe the advantages and disadvantages of each method with less than 100 words. (In other words, answer in up to 500 words in total which consists of (the advantages and disadvantages within 100 words) times 5 methods.)

問題番号 (Number): B-1

著作権上の理由で非表示

図は世界各地の森林流域において観測された年降水量と年流出水量の関係を示している。○は温帯の森林流域、●は熱帯の森林流域の観測値を示している。それぞれの観測値についての回帰直線の傾きと切片は、温帯で 0.98 と -679 mm、熱帯で 1.02 と -1,529 mm である。

- 1) 上図から得られる情報をもとに、温帯の森林流域における蒸発散量と降水量の地理的多様性の違いとその原因について論じなさい。
- 2) 上図から得られる情報をもとに、温帯の森林流域と熱帯の森林流域における蒸発散量の違いとその原因について論じなさい。



Not shown due to copyright

The figure shows the relationship between annual precipitation and annual discharge observed in forested catchments around the world. \circ indicates observed values in temperate forested catchments and \bullet indicates tropical forested catchments. The slopes and intercepts of the regression line for each observation are 0.98 and -679 mm in the temperate forests and 1.02 and -1,529 mm in the tropical forests.

- 1) Discuss the difference of the geographical diversity between the evapotranspiration and precipitation in the temperate forested catchment and its reasons based on the information obtained from the above figure.
- 2) Discuss the difference in evapotranspiration between the temperate forested catchment and the tropical forested catchment and its reasons based on the information obtained from the above figure.

問題番号 (Number): B-2

河川において 1,000km 以上の距離を移動する魚類の水平移動を音波テレメトリー用いてモニタリングする。この研究における音波テレメトリー技術の利点と問題点を論じなさい。

Suppose you try to monitor horizontal long-distance (more than 1,000 km) movements of fish in a river using acoustic telemetry. Discuss advantages and disadvantages of this monitoring technique for this study.

問題番号 (Number): B-3

以下の 5 題の小問から 3 題を選択し、解答しなさい。

Select three questions from the following five and answer them.

1. 最尤法と最小二乗法のそれぞれの特徴について説明せよ。

Explain characteristics of maximum likelihood method and least squares method.

2. NDVI(正規化植生指数)の原理と用途について説明せよ。

Explain principle and applications of NDVI (normalized difference vegetation index).

3. 生物学的種概念と系統学的種概念について説明せよ。

Explain the biological species concept and the phylogenetic species concept.

4. 窒素・炭素安定同位体比を用いた食性分析について説明せよ。

Explain the method for food habit analysis that uses nitrogen and carbon stable isotope ratios.

5. 森林生態系における窒素循環と炭素循環の構造的な違いを説明せよ。

Explain the difference of structures between nitrogen cycling and carbon cycling in forest ecosystems.

問題番号 (Number): B-4

ヒメウミガメ *Lepidochelys olivacea* は他のウミガメ類と同様に、メス成体が砂浜に上陸して産卵する。ヒメウミガメには、同じ個体群であっても単独で産卵をおこなう個体と集団で一斉に産卵をおこなう個体が存在する。この行動の多型が維持される仕組みとして、産卵巣密度に対して相対適応度が双峰型となる仮説（図）が提唱されている。

- 1) 適応度とは何か、説明しなさい。
- 2) 図のような相対適応度分布が生じる要因について、捕食と競合の効果をもとに論じなさい。なお、ここでは孵化幼体期を過ぎた個体の死亡率は同じであると仮定する。

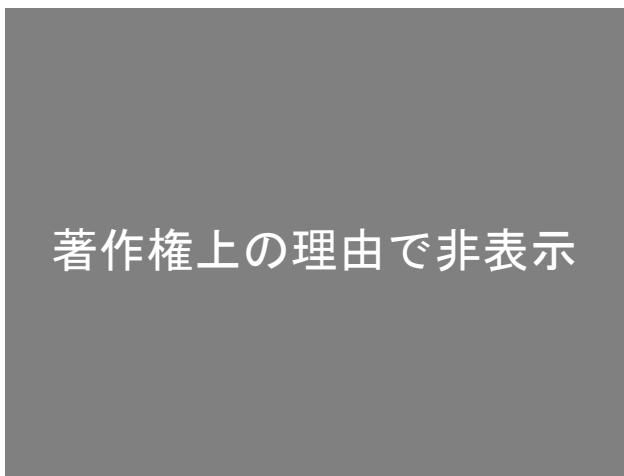


図 想定される産卵巣密度と相対適応度の関係

Source: Bernardo J, Plotkin PT (2007) An evolutionary perspective on the *arribada* phenomenon and reproductive behavioral polymorphism of olive ridley sea turtles (*Lepidochelys olivacea*). In: Plotkin PT (ed.), Biology and Conservation of Ridley Sea Turtles. The Johns Hopkins University Press, Baltimore, pp. 59–87.

Adult female olive ridley turtles *Lepidochelys olivacea* land on a beach for nesting as observed in the other sea turtle species. Even within a population, some females of olive ridley turtles nest solitarily, whereas some nest *en masse*. A hypothesis for maintenance of this behavioral plasticity assumes bimodal relative fitness in relation with nest density (Figure).

- 1) Explain what the fitness means.
- 2) Discuss the effects of predation and competition that result in the distribution of relative fitness in the Figure. The mortality of individuals is assumed to be same after hatchling period.



Figure The hypothesized relationship between nest density and relative fitness

Source: Bernardo J, Plotkin PT (2007) An evolutionary perspective on the *arribada* phenomenon and reproductive behavioral polymorphism of olive ridley sea turtles (*Lepidochelys olivacea*). In: Plotkin PT (ed.), Biology and Conservation of Ridley Sea Turtles. The Johns Hopkins University Press, Baltimore, pp. 59–87.

問題番号 (Number): D-1

図-1 は国際標準化機構（ISO）によって採用されたリスク管理のプロセスである。以下の問いに答えよ。

Figure-1 illustrates the risk management process adopted by the International Organization for Standardization (ISO31000). Answer the following questions.



Figure-1 Risk management process (ISO31000)

1. リスク管理の目的は何か。図-1 の各段階の行為の目的と方法、期待される結果を示すことで、このプロセスが何を達成することを目指すものかを示せ。
What are the goals of the risk management? By explaining the purpose, methods and outcomes of the actions in each of the stage in the Figure-1, describe what is expected to be achieved by use of the process.
2. このリスク管理プロセス(ISO31000)は災害リスク管理に適用可能か否か意見を述べよ。その理由を説明するとともに、必要ならその修正点を例示せよ。
State your opinion whether the ISO31000 risk management process can be applied for disaster risk management. Explain the reasons of your opinion and illustrate necessary revision if applicable.

問題番号 (Number): D-2

以下の問い合わせに答えよ

1. 次の数理計画問題について、以下の問い合わせに答えよ。

Answer the questions on the following mathematical programming problem:

$$\begin{aligned} & \max 2 \log x_1 + \log x_2 \\ & \text{subject to} \\ & \quad 2x_1 + x_2 \leq 8 \\ & \quad x_1 + ax_2 \leq 5 \\ & \quad x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

- (1) $a=1$ のとき、最適解を示せ。

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

- (2) 直線 $2x_1 + x_2 = 8$ 直線 $x_1 + ax_2 = 5$ との交点を X_0 とし、 X_0 が最適解となる a の範囲を示せ。

Let X_0 be the intersection point of lines $2x_1 + x_2 = 8$ and $x_1 + ax_2 = 5$.

Show the interval of a in which the optimal solution of the problem is X_0 .

-

問題番号 (Number): D-3

減災研究における「アクションリサーチ」、「実践科学」とは何か。これら2つの概念について具体的な事例をまじえて論じなさい。

What are “action research” and “implementation science” in disaster reduction studies?
Discuss these two concepts using concrete examples.

問題番号 (Number): D-4

以下の問い合わせに答えよ。

- (1) Web 経由で地図データを提供するための以下の国際規格について説明しなさい。
 - (a) Web Map Service (WMS)
 - (b) Web Feature Service (WFS)
 - (c) Web Map Tile Service (WMTS)
- (2) GIS における以下の数値標高モデルの原理、利点、欠点について説明しなさい
 - (a) Digital Surface Model (DSM)
 - (b) Digital Terrain Model (DTM)
 - (c) Triangulated Irregular Network (TIN)

Answer the following questions.

- (1) Explain the international standards for serving georeferenced map data through the Web as follows;
 - (a) Web Map Service (WMS)
 - (b) Web Feature Service (WFS)
 - (c) Web Map Tile Service (WMTS)
 - (2) Describe a fundamental principle, advantages, and disadvantages of the Digital Elevation Models in GIS as follows;
 - (a) Digital Surface Model (DSM)
 - (b) Digital Terrain Model (DTM)
 - (c) Triangulated Irregular Network (TIN)
-
-

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）では、15の緊急支援機能（ESF；Emergency Support Function）が規定されている。このうち、以下の2つの機能について「情報」の観点から具体例に基づき論じなさい。

- (a) Emergency Support Function #7 – Logistics Management and Resource Support
- (b) Emergency Support Function #8 – Public Health and Medical Services

The National Emergency Management System (NIMS), a US crisis management standard, defines 15 Emergency Support Functions (ESFs). Discuss these two functions from the viewpoints of “information” with a concrete example as follows;

- (a) Emergency Support Function #7 – Logistics Management and Resource Support
- (b) Emergency Support Function #8 – Public Health and Medical Services

問題番号 (Number): M-1

- (1) 遠隔医療とは何か簡潔に述べよ
(2) 遠隔医療の意義の一部として、「医療格差の解消」「医療の効率化」が挙げられている。前問の解答を踏まえ、遠隔医療がそれらに寄与する理由を述べよ

- (1) What is telemedicine? Explain concisely.
(2) Telemedicine is expected to contribute to “resolving disparity in medicine” and “effective healthcare delivery”. Considering your answer above, explain how telemedicine contribute to them.
-
-

問題番号 (Number): M-2

医療情報の電子化が進むにつれて、今まで組織ごとに保有してきたデータを集約して情報共有基盤を構築しようとする動きが活発化している。しかし、情報共有基盤を構築する際には克服しなければならない技術的課題がいくつかある。

- (1) 医療情報の共有基盤を構築することにより得られる臨床的利点および学術的利点について述べよ。
(2) 医療情報の共有基盤を構築する際の技術的課題に関する具体例を1つ挙げ、その解決方法について述べよ。

The progress of digitization of healthcare data encourages the development of healthcare data sharing systems, which are to aggregate electronic healthcare data stored in various organization and to share them. However, there are some technical challenges to realize such systems.

- (1) Describe a clinical advantage and an academic advantage, which such systems provide.
(2) Describe one of the technical challenges and a plausible solution for it.

問題番号 (Number): M-3

CT (Computed Tomography) 検査や MRI (Magnetic Resonance Imaging) 検査, PET (Positron Emission Tomography) 検査, 超音波検査など, 画像を用いた検査にはいくつのかの種類がある. 加えて、複数の画像を融合（フェージョン）することで双方の利点を反映した画像に基づく検査も可能になっている. しかし、融合画像を診療に用いるには、様々な工夫が必要である.

- (1) フュージョンの一例を示し, その融合による利点について説明せよ.
- (2) 施されなければならない工夫を一つ挙げ, 解決策を提案せよ.

Various body scanning technologies, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasonography, and so on, are widely used in clinical field. Nowadays, “fusion images”, which are generated by merging multiple images, are also applied to clinical activities. However, there are several challenges to apply them.

- (1) Give an example of fusion images, and explain the merits obtained by the fusion.
 - (2) Describe one of the challenges and a plausible solution for it.
-
-

問題番号 (Number): M-4

病院内で車椅子を貸し出すシステムがあると仮定する。貸し出しあはウェブサイトで行われ、利用者は先着順に借りることができる。備品が全て貸し出されていたら、貸し出し希望者は、返却されるのを待たなければならぬ。車椅子が 10 台あると仮定し、当該システムのフローチャートを示せ。

Suppose that there is a system for lending wheelchairs. Applicants can apply via a web site, and the wheelchairs are lent out under first-come-first-served bases. When all the equipment has been lent, the applicant has to wait for one to be returned. Assume there are 10 wheelchairs to be lent, and draw a flow chart of the system.

問題番号 (Number): M-5

ある特定の情報を含む診療文書を抽出する装置を、機械学習を用いて生成することを考える。特定の情報を含む診療文書 100 文書と特定の情報を含まない診療文書 100 文書を学習データとして装置を生成し、生成された装置の性能（精度と再現率）を交差検定によって測定した。さて、この装置を用いてある病院の保有する全診療文書から特定の情報を含む診療文章を抽出した場合、上記で計測した性能となると期待されるか考察せよ。

Assume extracting clinical documents including some specific information using a classifier developed via machine learning technique.

The classifier was developed using a training data set, which consists of 100 documents with the specific information and 100 documents without the specific information. The performance (precision and recall) of the developed classifier was measured using cross validation.

Discuss whether we can expect the same performance as measured above, when we apply the developed classifier to extract documents with the specific information from all documents in a certain hospital.