

平成 27 年度 10 月期入学 / 平成 28 年度 4 月期入学
京都大学 大学院情報学研究科
修士課程 知能情報学専攻 入学者選抜試験問題
(専門科目)

平成 27 年 8 月 6 日 9:00～12:00

【注意】

1. 問題冊子はこの表紙を含めて 33 枚ある。
2. 試験開始の合図があるまで中を見てはいけない。
3. 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。
4. 問題は下記 16 題であり、日本語と英語の両方で出題されている。このうちいずれか 4 題を選択し、解答しなさい。

生命情報学(問題番号:B-1～B-4)	1-8 ページ
心理学(問題番号:P-1～P-4)	9-16 ページ
計算機科学, 電気電子工学(問題番号:T-1～T-8)	17-32 ページ

5. 特に指定のない限り、日本語または英語で解答すること。
6. 解答用紙に記載されている注意事項についても留意すること。

*The Japanese version of this document is the prevailing and authoritative version;
the English translation below is provided for reference only*

October 2015 Admissions / April 2016 Admissions
Entrance Examination for Master's Program
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
(Specialized Subjects)

August 6, 2015
9:00 - 12:00

NOTES

1. This is the Question Booklet in 33 pages including this front cover.
2. Do not open the booklet until you are instructed to start.
3. After the exam has started, check the number of pages and notify proctors (professors) immediately if you find missing pages or unclear printings.
4. There are 16 questions, written in Japanese and English. The questions are classified as listed below. **Choose and answer 4 questions.**

Bioinformatics (Question Numbers B-1 to B-4)	Pages 1 to 8
Psychology (Question Numbers P-1 to P-4)	Pages 9 to 16
Computer Science, Electrical and Electronic Engineering (Question Numbers T-1 to T-8)	Pages 17 to 32

5. Write your answer in Japanese or English, unless otherwise specified.
6. Read carefully the notes on the Answer Sheets as well.

遺伝子の転写に関する以下の設問に答えよ。

設問1 転写調節を行うタンパク質には、DNA 結合モチーフと転写活性化モチーフがよくみられる。

- (1) 転写調節を行うタンパク質にみられる DNA 結合モチーフをひとつあげ、その特徴について簡潔に説明せよ。
- (2) DNA 結合モチーフが DNA に結合できる仕組みについて簡潔に説明せよ。

設問2 転写因子 X の転写活性について解析する目的で、転写因子 X の欠失変異体タンパク質(m1 - m3)を作製した (図 1)。これらの欠失変異体タンパク質と転写因子 X の結合配列を用いてゲルシフトアッセイを行った (図 2)。

- (1) ゲルシフトアッセイの方法の概要を説明せよ。
- (2) 図 2 でレーン 1 - 5 に共通して見られる下のバンドはなにか、説明せよ。
- (3) 図 2 のレーン 1, 2 では 1 本のバンドが観察され、レーン 3, 4 では 2 本のバンドが観察された。レーン 3, 4 で 2 本のバンドが観察される理由について説明せよ。
- (4) 図 2 のレーン 3, 4 において 2 本のバンドのうち上のバンドの移動度が異なる理由について説明せよ。
- (5) 図 2 のレーン 5 で 4 本のバンドが観察される理由について説明せよ。



図 1 作成した欠失変異体の模式図。

N : N 末端。C : C 末端。

m1	-	+	-	-	-
m2	-	-	+	-	+
m3	-	-	-	+	+

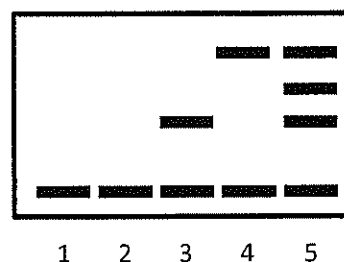


図 2 ゲルシフトアッセイ。上の表はゲルシフトアッセイで用いた欠失変異体を示す。

Question is translated in English in the section below; this translation is given for reference only.

Answer the following questions about gene transcription.

Q.1 In general, transcription factors have a DNA-binding motif and a transcription activating motif.

- (1) Name the DNA-binding motif found in transcription factors and describe its features briefly.
- (2) Explain briefly how the DNA-binding motif binds to DNA.

Q.2 To analyze transcriptional activities of a transcription factor X, deletion mutants of the transcription factor X (m1 - m3) were constructed (Fig. 1). Using these mutants and DNA-binding sequence of the transcription factor X, gel-shift assay was performed (Fig. 2).

- (1) Summarize the gel-shift assay method.
- (2) What is the lowest band found in lanes 1 - 5 in Fig. 2?
- (3) In Fig. 2, a single band was observed in lanes 1 and 2, and two bands were observed in lanes 3 and 4. Explain why two bands were observed in lanes 3 and 4.
- (4) In Fig. 2, explain why the upper bands in lanes 3 and 4 have different mobilities.
- (5) In Fig. 2, explain why four bands were observed in lane 5.

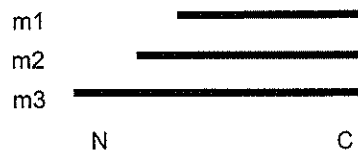


Fig.1 Schematic representation of deletion mutants of the transcription factor X.

N: N-terminus. C: C-terminus.

m1	-	+	-	-	-
m2	-	-	+	-	+
m3	-	-	-	+	+

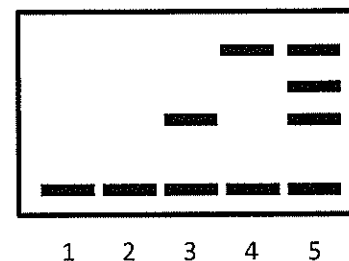


Fig. 2 Gel-shift assay. The upper part of this figure shows which mutants were used in this assay.

タンパク質に関して以下の設問に答えよ。

設問1 タンパク質についての以下の語句を簡潔に説明せよ。

- (1) α ヘリックス
- (2) コイルドコイル
- (3) 等電点
- (4) 四次構造
- (5) タンパク質ドメイン

設問2 あるタンパク質を SDS ポリアクリルアミド電気泳動法で解析した。タンパク質サンプルを還元剤で処理しない場合は 1 本のバンドが検出された。同じタンパク質の別のサンプルを還元剤で処理した場合も 1 本のバンドが検出されたが、還元剤で処理しない場合に比べて移動度が異なっていた。なぜ還元剤処理によってこの違いが生まれたかを説明せよ。

設問3 サブユニットの結合に関与する弱い相互作用（非共有結合）をひとつ挙げ、説明せよ。

設問4 あるタンパク質溶液を加熱するとタンパク質が不溶化した。このタンパク質溶液を還元剤と界面活性剤とともに加熱すると不溶化したタンパク質が可溶化する。

- (1) 加熱すると不溶化する理由について説明せよ。
- (2) 還元剤と界面活性剤が可溶化に必要な理由について説明せよ。

Question is translated in English in the section below; this translation is given for reference only.

Answer the following questions about proteins.

Q.1 Explain briefly the following words about proteins.

- (1) α -helix
- (2) Coiled coil
- (3) Isoelectric points
- (4) Quaternary structure
- (5) Protein domain

Q.2 A protein is analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). One sample of this protein, which was not treated with reducing agents, shows a single band on the SDS-PAGE gel. Another sample of the same protein, which was treated with these reducing agents, shows also a single band but with a different mobility than the non-treated sample. Explain why these reducing agents make this difference.

Q.3 Describe in detail one of the weak interactions (non-covalent bonding) that play a role in subunit assembly.

Q.4 When a protein solution is heated, proteins become insoluble. These insoluble proteins can be dissolved by heating them in a solution containing both a detergent and a reducing agent.

- (1) Explain why heating soluble proteins causes them to become insoluble.
- (2) Explain why both a reducing agent and a detergent are required to dissolve insoluble proteins.

生命情報学に関する以下の語句の中から8つを選んで、それぞれ4行以上10行以内で説明せよ。必要に応じて図を用いてもよい。

1. トランスポゾン
2. 遺伝暗号表
3. β シート
4. モルテングロビュール状態
5. GenBank データベース
6. Pfam データベース
7. 近隣結合法
8. ローカルアラインメント
9. フラグメントアセンブリ法
10. 代謝流束解析

Question is translated in English in the section below; this translation is given for reference only.

Choose and explain 8 items from the following 10 items related to bioinformatics, where 4–10 lines must be used for each explanation, and figure(s) can be used if necessary.

1. Transposon
2. Genetic code table
3. β -sheet
4. Molten globule state
5. GenBank database
6. Pfam database
7. Neighbor-joining method
8. Local alignment
9. Fragment assembly method
10. Metabolic flux analysis

長さ n の DNA 配列 x と y に対して, その距離 $d(x, y)$ を以下で定義する.

$$d(x, y) = |\{i \mid x[i] \neq y[i], i = 1, \dots, n\}|$$

ただし, 配列 (文字列) s に対し $s[i]$ は i 番目の文字を表すものとする. よって, $x = \text{ACGT}$, $y = \text{AAGT}$, $z = \text{GACT}$ の時, $d(x, y) = 1$, $d(x, z) = 3$, $d(y, z) = 2$ となる. 以下の設問に答えよ.

設問 1 $x = \text{ACGTAT}$, $y = \text{ACTGCT}$ の時, $d(x, z) = 1$, $d(y, z) = 2$ を満たす DNA 配列 z をすべて求めよ.

設問 2 長さ n の任意の DNA 配列 x に対して, $|\{y \mid d(x, y) = d_0\}|$ を与える d_0 と n に関する簡潔な式を理由とともに示せ. ただし, d_0 は n 以下の非負整数とし, 任意の非負整数 m に対し $\binom{m}{0} = 1$ とする.

設問 3 長さ n の DNA 配列 x, y があり, すべての $i = 1, \dots, n$ について $x[i] \neq y[i]$ が成立するものとする. この時, $d(x, z) = d_1$, $d(y, z) = d_2$ を満たす DNA 配列 z が存在するための必要十分条件は以下の式で与えられる (ただし, d_1, d_2 は n 以下の非負整数).

$$n \leq \boxed{} \leq \boxed{}$$

空欄を埋めた式を理由とともに示せ.

設問 4 設問 3 と同様, 長さ n の DNA 配列 x, y があり, すべての $i = 1, \dots, n$ について $x[i] \neq y[i]$ が成立するものとする. 設問 3 の条件が成立する時, $d(x, z) = d_1$, $d(y, z) = d_2$ を満たす DNA 配列 z の個数は以下の式で与えられる.

$$\boxed{} \cdot \binom{\boxed{}}{d_2} \cdot \binom{d_2}{\boxed{}}$$

空欄を埋めた式を理由とともに示せ.

Question is translated in English in the section below; this translation is given for reference only.

For DNA sequences x and y of length n , we define a distance $d(x, y)$ between x and y by

$$d(x, y) = |\{i \mid x[i] \neq y[i], i = 1, \dots, n\}|,$$

where $s[i]$ denotes the i -th letter of a sequence (string) s . For example, $d(x, y) = 1$, $d(x, z) = 3$, and $d(y, z) = 2$ hold for $x = \text{ACGT}$, $y = \text{AAGT}$, and $z = \text{GACT}$.

Answer the following questions.

Q.1 List all DNA sequences z which satisfy $d(x, z) = 1$ and $d(y, z) = 2$, where $x = \text{ACGTAT}$ and $y = \text{ACTGCT}$.

Q.2 Describe a simple mathematical formula on d_0 and n that gives $|\{y \mid d(x, y) = d_0\}|$ and explain its derivation, where x is a DNA sequence of length n , and d_0 is an integer with $0 \leq d_0 \leq n$. Use the convention $\binom{m}{0} = 1$ for any non-negative integer m .

Q.3 Let x and y be DNA sequences of length n such that $x[i] \neq y[i]$ holds for all $i = 1, \dots, n$. The following inequalities give a necessary and sufficient condition for the existence of a DNA sequence z satisfying $d(x, z) = d_1$ and $d(y, z) = d_2$, where d_1 and d_2 are non-negative integers not exceeding n :

$$n \leq \boxed{} \leq \boxed{}.$$

Complete the formula by filling the blank boxes, and explain its derivation.

Q.4 As in Q.3, let x and y be DNA sequences of length n such that $x[i] \neq y[i]$ holds for all $i = 1, \dots, n$. If the condition given in Q.3 is satisfied, the number of DNA sequences z satisfying $d(x, z) = d_1$ and $d(y, z) = d_2$ is given by

$$\boxed{} \cdot \binom{\boxed{}}{d_2} \cdot \binom{d_2}{\boxed{}}.$$

Complete the formula by filling the blank boxes, and explain its derivation.

設問 大脳視覚野に関連する以下の概念について簡潔に説明せよ。数式や図を用いてもよい。

- (1) 受容野 (Receptive field)
- (2) 逆相関法 (Reverse correlation)
- (3) 皮質拡大係数 (Cortical magnification factor)
- (4) ガボール関数 (Gabor function)
- (5) エネルギーモデル (Energy model)
- (6) 非古典的受容野 (Non-classical receptive field)
- (7) スパース性 (Sparseness)
- (8) 理想的観察者 (Ideal observer)

Question is translated in English in the section below; this translation is given for reference only.

Q. Give a brief explanation on each of the following items in the context of visual cortex. Mathematical expressions and/or figures may be used.

- (1) Receptive field
- (2) Reverse correlation
- (3) Cortical magnification factor
- (4) Gabor function
- (5) Energy model
- (6) Non-classical receptive field
- (7) Sparseness
- (8) Ideal observer

ある実験で、Figure 1 の 16 種類の刺激について、文字の間隔を 3 段階に変化させた計 48 種類の画面を繰り返しランダム順に提示した。実験参加者の課題は、中央の文字が H か K の場合には右手に割り当てられたキーを、C か S の場合には左手に割り当てられたキーをできるだけ素早く、かつ正確に押すことであった。刺激画面が表示されてから、キーが押されるまでの反応時間を測定した。Figure 2 は、平均反応時間を刺激と反応の組み合わせ条件（Figure 1 参照）別にプロットしたものである。

以下の設問に答えよ。

設問 1 それぞれのキーに 2 つずつ文字が割り当てられているのはなぜか説明せよ。

設問 2 noise response incompatible 条件の反応時間が noise response compatible 条件の反応時間よりも遅いのはなぜか説明せよ。

設問 3 noise response incompatible 条件の反応時間と noise response compatible 条件の反応時間の差が、between-letter spacing が大きくなるに伴って、減少するのはなぜか説明せよ。

Noise same as target	Noise response compatible	Noise response incompatible	
HHHHH	KKHKK	CCHCC	SSHSS
KKKKK	HHKHH	CCKCC	SSKSS
CCCCC	SSCSS	HHCHH	KKCKK
SSSSS	CCSCC	HHS HH	KKSKK

Figure 1. Stimulus displays used in the experiment: Four displays for the “noise same as target” condition, four displays for the “noise response compatible” condition, and eight displays for the “noise response incompatible” condition.

Non-Disclosure

(Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, 16, 143–149.を改変)

Question is translated in English in the section below; this translation is given for reference only.

In an experiment, sixteen sets of letters (shown in Figure 1) were repeatedly presented with three different between-letter spacings in a random order. The participants' task was to press a key assigned to their right hand if the central letter was H or K, and to press a key assigned to their left hand if the central letter was C or S, as soon as possible while minimizing errors. Reaction time from the onset of a stimulus display to the initiation of a key response was measured. Figure 2 shows mean reaction times as a function of these between-letter spacings, separately plotted for three experimental conditions (see Figure 1).

Answer the following questions.

Q.1 Why were two letters assigned to each key?

Q.2 Why were reaction times for the "noise response incompatible" condition slower than those for the "noise response compatible" condition?

Q.3 Why did the difference of reaction times between the "noise response compatible" condition and the "noise response incompatible" condition decrease as the between-letter spacing increases?

Noise same as target	Noise response compatible	Noise response incompatible	
HHHHH	KKHKK	CCHCC	SSHSS
KKKKK	HHKHH	CCKCC	SSKSS
CCCCC	SSCSS	HHCHH	KKCKK
SSSSS	CCSCC	HHS HH	KKSKK

Figure 1. Stimulus displays used in the experiment: Four displays for the "noise same as target" condition, four displays for the "noise response compatible" condition, and eight displays for the "noise response incompatible" condition.

Non-Disclosure

Modified from Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, 16, 143–149.

図 1 は、エビングハウス錯視の 2 つの条件 (A, B) を図示している。各パネルには、周囲を他の円 (大きい円あるいは小さい円) に囲まれた 2 つのターゲット円が描かれている。また各パネルの下にも、ターゲット円と同じ円が示されている。条件 A では、2 つのターゲット円は物理的に同じであるが、一般に、大きさは異なって知覚される。一方、条件 B では、ターゲット円の物理的な大きさを変えることで、知覚される大きさをほぼ同じにしている。この錯視を用いた実験について、以下の設問に答えよ。



図 1, 2: Aglioti et al. (1995) *Current Biology* 5: 679-685 を改変

設問 1 実験参加者の親指と人差し指にセンサーをつけ、ターゲット円をつまむときの指の開きを調べた。図 2 は、B の条件 (知覚は同じだが物理的大きさが異なる) で、物理的に大きいほうの円 (実線) と小さいほうの円 (破線) をつまむときの、人差し指と親指の間隔の変化をプロットしたものである。この結果からわかることを、「行為のための視覚」 (Vision for action) という用語を用いて説明せよ。

設問 2 次に、右利きと左利きの実験参加者を新たに集めて、それぞれに左手と右手で設問 1 と同じ課題を遂行させた。図 3 は、A の条件 (知覚は異なるが物理的大きさは同じ) の結果を示す。まず各被験者について、A の条件でターゲット円をつまむときの人差し指と親指の間隔の最大値を求めた。そして、大きく知覚される円をつまむときと、小さく知覚される円をつまむときの、最大値の差を計算した。図 3 の棒グラフは、その値 (最大値の差) を示している。縦軸の値が大きいほど、指の開きが錯視に影響される程度が大きいことを示す。図 3 のデータに関する統計検定の結果は以下の通りであった。

“A two-way ANOVA between group (right- and left-handers) and hand (right and left hand) revealed a main effect of hand [$F(1, 38) = 24.86, p < 0.001$], no effect of group [$F(1, 38) = 0.20, p > 0.05$], and no significant interaction between group and hand [$F(1, 38) = 0.10, p > 0.05$].”

- (1) この実験の要因計画を推定して説明せよ。
- (2) 結果の概要を説明し、脳の左右差の観点から考察せよ。

Question is translated in English in the section below; this translation is given for reference only.

Figure 1 illustrates Condition A and Condition B of the Ebbinghaus illusion. The panel for each condition displays two target circles surrounded by the other (smaller and larger) circles. These two target circles are displayed separately at the bottom of each panel. In Condition A, participants usually report that the target circles are different in size, despite the fact that the two circles are physically identical. By contrast, in Condition B, the two circles are different in actual size, when the two target circles appear identical. Answer the following questions related to experiments that used this illusion.

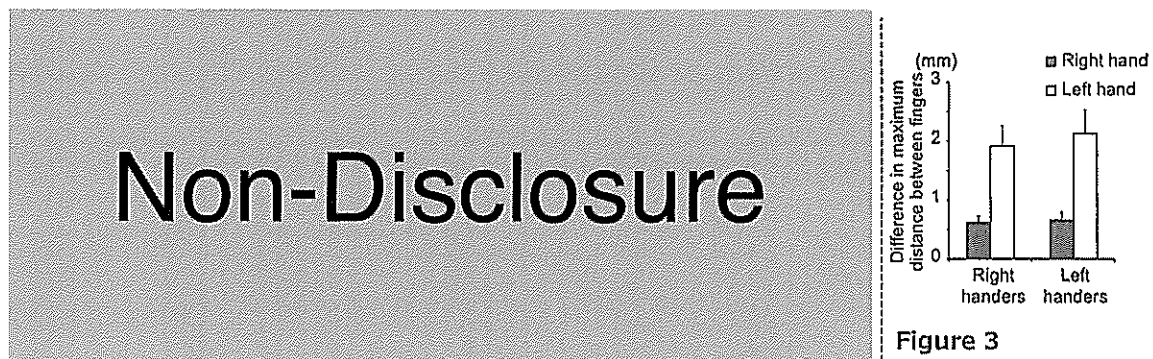


Figure 3

Figures 1 and 2: modified from Aglioti et al. (1995) *Current Biology* 5: 679-685

Q.1 Sensors were attached to the index finger and thumb of a participant, and the distance between them was recorded while the participant picked up the central target circles. In Figure 2, the distance between the index finger and thumb is plotted as a function of time, as the participant was asked to pick up the larger (solid curve) and smaller (dotted curve) circles under Condition B (perceptually identical, but physically different). Explain this result using the concept of “vision for action.”

Q.2 Two groups of participants (right-handers and left-handers) carried out the same task as the one in Q.1 using either their right or left hand. Figure 3 shows the result in Condition A (perceptually different but physically identical). For each participant, the experimenters extracted the maximum distance between the index finger and thumb during the task of picking up circles under Condition A. The maximum distance when picking up the circle which is perceived as smaller is subtracted from that observed when picking up the circle perceived as larger. In Figure 3, the bar chart of that difference is displayed for both right and left handers, using either hand. Larger values in that difference indicate that the illusion has a stronger impact. The statistical result for the data in Figure 3 was stated as follows. “A two-way ANOVA between group (right- and left-handers) and hand (right and left hand) revealed a main effect of hand [$F(1, 38) = 24.86, p < 0.001$], no effect of group [$F(1, 38) = 0.20, p > 0.05$], and no significant interaction between group and hand [$F(1, 38) = 0.10, p > 0.05$].”

- (1) Explain the factorial design of this experiment.
- (2) Summarize the result, and discuss it in terms of the laterality of the brain.

音韻作業記憶の研究を目的として単語の復唱実験を行った。1 音節の英単語 “sum” や “hate”などを連続的に 5 単語提示する条件と、5 音節の英単語 “association” や “representative”などを連続的に 5 単語提示する条件を実施したときに、実験参加者には英単語の復唱を要求した。そのときの復唱した英単語の正答率に関する実験結果を、統制条件として図 1 に示す。また、統制条件と同じ刺激提示の後に、ある課題を実験参加者に要求することで、図 1 の課題条件に示す英単語復唱の正答率を得た。

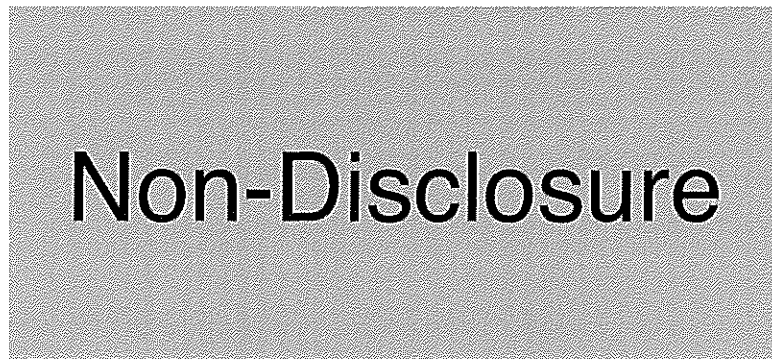


図 1

Baddeley, Thomson & Buchanan (1975)

Journal of Verbal Learning and Verbal Behavior 14, 575-589 の Fig. 5 を一部改変

以下の設問に答えよ。

設問 1 図 1 に示す①と②は 1 音節単語提示条件、5 音節単語提示条件のいずれであると考えられるか。その理由とともに述べよ。

設問 2 図 1 の課題条件は音韻ループを妨害することを目的としてある課題を実験参加者に要求したときの結果である。どのような課題を行ったと考えられるかを述べよ。

設問 3 課題条件において音韻ループを妨害することで、なぜ単語復唱の正答率が低下するとともに、語長効果が消失すると考えられるかを述べよ。

設問 4 Baddeley の作業記憶モデルでは、音韻ループの他にも 2 種類の機能要素があると提案している。この 2 種類の要素と、その機能を述べよ。

Question is translated in English in the section below; this translation is given for reference only.

A word recall experiment was performed to investigate the phonological working memory. The experiment constituted of one-syllable and five-syllable conditions. The one-syllable condition was to present five one-syllable words, e.g., “sum” and “hate,” sequentially. The five-syllable condition was to present five five-syllable words, e.g., “association” and “representative.” Participants were required to recall the words verbally just after the presentation. Figure 1 shows the accuracy for participants’ recall with the label of *Control*. Figure 1 also includes the accuracy for participants’ recall, where the participants were required to carry out an additional task just after the presentation, with the label of *Task*.

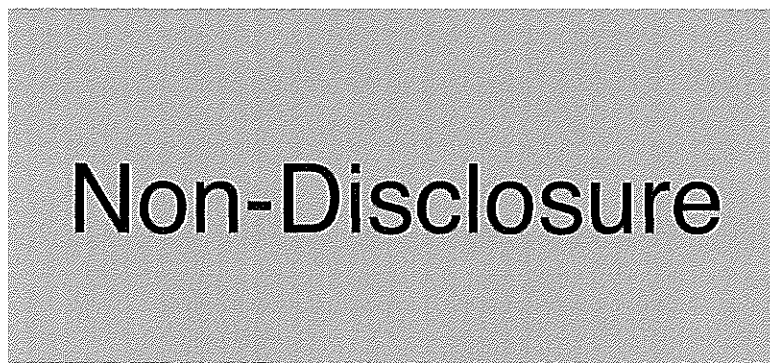


Figure 1

Modified from Fig. 5 in Baddeley, Thomson & Buchanan (1975), *Journal of Verbal Learning and Verbal Behavior* 14, 575-589

Answer the following questions.

Q.1 Identify the one-syllable condition and five-syllable condition accuracy lines in Fig. 1. Provide a detailed answer.

Q.2 The accuracies with the label of *Task* in Fig. 1 were acquired when the participants were required to carry out an additional task to distract the phonological loop. Give an example of such a distracting task.

Q.3 Discuss why the accuracy for participants’ recall dropped and the word length effect disappeared, when the phonological loop was distracted in the *Task* condition.

Q.4 Baddeley proposed a working memory model, in which there were two other functional components in addition to the phonological loop. Name these two components and describe their functions.

X は1, -1のいずれかの値をそれぞれ確率 $\Pr(X = 1) = \frac{2}{3}$, $\Pr(X = -1) = \frac{1}{3}$ でとる確率変数である。 X をあるアナログ通信路を介して送信すると、受信者側では $Y = X + Z$ が観測される。ただし、 Z は確率密度関数 $f(z) = \frac{1}{2}\exp(-|z|)$ に従う確率変数とする。受信者は Y の値 y から X の値 x を推定する。受信者はある定数の閾値 t を用いて、 $y \geq t$ ならば $x = 1$ 、 $y < t$ ならば $x = -1$ と復号するものとする。

設問 1 X のエントロピーを求めよ。

設問 2 $t = 0$ とする。 $x = 1$ であったとき、これを受信者が誤って-1と復号する確率を求めよ。

設問 3 復号誤り率が最も小さくなるような t を求めよ。

Question is translated in English in the section below; this translation is given for reference only.

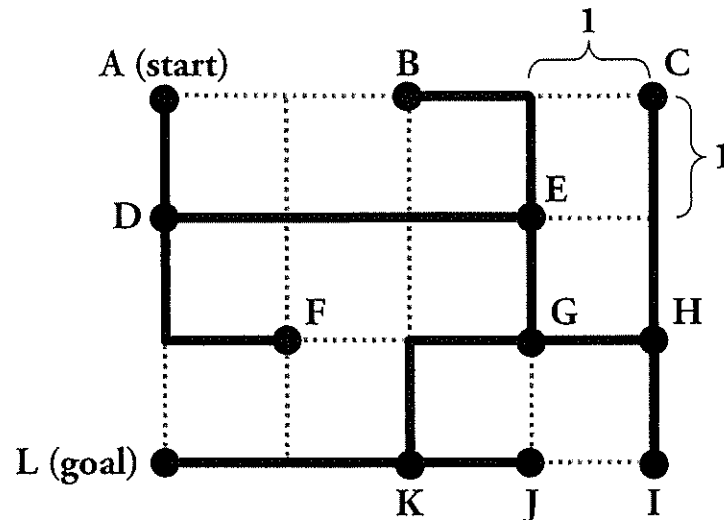
X is a random variable that takes the value 1 with probability $\Pr(X = 1) = \frac{2}{3}$ or -1 with $\Pr(X = -1) = \frac{1}{3}$. When we transmit X over an analog channel, a receiver receives $Y = X + Z$ at the other end, where a random variable Z follows a probability density function $f(z) = \frac{1}{2} \exp(-|z|)$. The goal of the receiver is to guess the value x taken by X given the observed value y taken by Y . Given a threshold t , the receiver decides that $x = 1$ if $y \geq t$ or $x = -1$ if $y < t$.

Q.1 Compute the entropy of X .

Q.2 Assume $t = 0$. Find the probability that the receiver makes the wrong guess that $x = -1$ when in fact x was equal to 1.

Q.3 Find t that gives the smallest decoding error rate.

下図は、縦横の間隔がともに1であるグリッド（点線）上に定義された迷路（実線）を示している。この迷路は、AからLまでの12個の節点およびそれら節点を結ぶ11個の通路からなる。いま、計算機を用いて、逐次的に節点を展開しながら、AからLまでの経路を探索する問題を考える。ただし、計算機には、すべての節点と通路の位置が事前に登録されており、評価関数の最良値が同じである節点が複数存在する場合はアルファベット順に展開するものとする。このとき、以下の設問に答えよ。



設問1 この問題の状態空間を図示せよ。

設問2 幅優先探索を用いてこの問題を解け。ただし、探索の進行に伴うオープンリスト（未展開の節点の順序付き集合）の逐次的な変化を示すこと。表記例：[A] → [D] → [B, E]

設問3 一般に、深さが d 、最大分岐数が b である木を探索する場合の、幅優先探索の最悪空間計算量および最悪時間計算量について説明せよ。

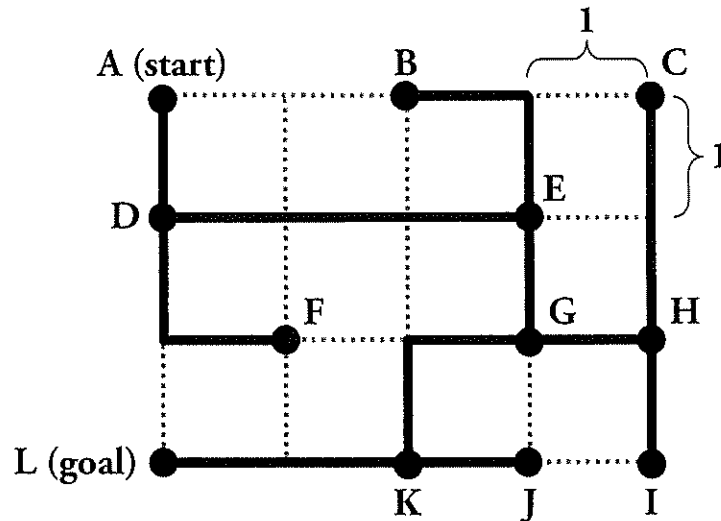
設問4 ダイクストラ法を用いてこの問題を解け。ただし、次に展開する節点を選ぶための評価関数を定義し、探索の進行に伴うオープンリストの逐次的な変化を評価関数の値とともに示すこと。表記例：[A(3)] → [D(8)] → [B(4), E(6)]

設問5 A*アルゴリズムを用いてこの問題を解け。ただし、設問4と同様の形式で答えること。

設問6 一般に、次に展開する節点の選び方の観点から、幅優先探索、ダイクストラ法およびA*アルゴリズムの関係を説明し、各アルゴリズムの完全性と最適性について述べよ。

Question is translated in English in the section below; this translation is given for reference only.

The figure shown below illustrates a maze (solid lines) defined on a unit grid (dotted lines) with an interval of 1. This maze consists of 12 nodes labeled from A to L and 11 edges connecting those nodes. The problem is to search for a path from A to L by using a computer while incrementally expanding the nodes. Suppose that the positions of all the nodes and edges are registered on the computer in advance and that the next-expanded node is selected among those with the best evaluation score using the alphabetical order. Answer the following questions.



- Q.1 Illustrate the state space of the above-defined maze problem.
- Q.2 Solve the above-defined maze problem using the breadth-first-search algorithm. Show how the open list (the ordered list of unexpanded nodes) is changed according to the search progression, *e.g.*, $[A] \rightarrow [D] \rightarrow [B, E]$.
- Q.3 Explain the worst-case space and time complexities of the breadth-first-search algorithm when it is used for searching a tree with a depth of d and with a maximum branching factor of b .
- Q.4 Solve the above-defined maze problem using Dijkstra's algorithm. Define an evaluation function for selecting the next-expanded node, and show how the open list is changed according to the search progression. Each node in the open list should be given an evaluation score, *e.g.*, $[A(3)] \rightarrow [D(8)] \rightarrow [B(4), E(6)]$.
- Q.5 Solve the above-defined maze problem using the A* algorithm. Answer this question in the same way as Q.4.
- Q.6 From a general point of view, explain the relationships between the breadth-first-search algorithm, Dijkstra's algorithm, and the A* algorithm in terms of how to select the next-expanded node, and then explain the completeness and optimality of each algorithm.

d 次元のパターンベクトル $X = (x_1, \dots, x_d)$ に対して、 $g(X) = w_1x_1 + \dots + w_dx_d + w_{d+1}$ で表される識別関数が正のときクラス 1、負のときクラス 0 に分類・識別する。

設問1. 誤り訂正学習法 (パーセプトロン学習法) に基づいて、重み $W = (w_1, \dots, w_d, w_{d+1})$ を学習する方法の概要を述べよ。またこの学習法により解が見つかる条件を述べよ。

設問2. $f(z)$ をシグモイド関数とし、 $f(g(X))$ を新たに識別関数として定義する。その値域とそのうちクラス 1 と判定される値の範囲を述べよ。

設問3. X がクラス 1 のとき $t = 1$ 、クラス 0 のとき $t = 0$ とする教師信号 t に対する二乗誤差

$J = \frac{1}{2}(t - f(g(X)))^2$ を定義する。このとき、 $\frac{\partial J}{\partial w_i}$ を求めよ。ただし、 $f(z)$ の導関数は

$f'(z) = f(z)(1 - f(z))$ で与えられる。

誤差逆伝播学習法は、重み w_i を $-c \frac{\partial J}{\partial w_i}$ (ただし c は定数) に応じて更新するものである。設問 1 の誤り訂正学習法がこれを簡潔にしたものであることを式の上で説明せよ。

設問4. 効果的なパターン分類・識別を実現するには、識別関数を多段にした多層ネットワークを構成する必要がある。その理由を述べよ。また 2 段のネットワークを構成する場合、上記の $g(X)$ を複数 $(g_1(X), \dots, g_k(X))$ 用意して、その識別結果の重み付き多数決をとることに相当することを示せ。

Question is translated in English in the section below; this translation is given for reference only.

For a d -dimensional pattern vector $X = (x_1, \dots, x_d)$, a discriminant function $g(X) = w_1x_1 + \dots + w_dx_d + w_{d+1}$ classifies X into class 1 when $g(X)$ is positive or class 0 when negative.

Q.1 Describe a procedure of training weights $W = (w_1, \dots, w_d, w_{d+1})$ based on the error correction (perceptron) algorithm. Describe also conditions under which the algorithm reaches a solution.

Q.2 Let f be the sigmoid function and $f(g(X))$ be a re-defined discriminant function. Describe the range of f . For which values of $f(g(X))$ is X classified into class 1?

Q.3 Given a training label t ($t = 1$ for class 1 and $t = 0$ for class 0), a square error $J = \frac{1}{2}(t - f(g(X)))^2$ is defined. Calculate the partial derivative $\frac{\partial J}{\partial w_i}$. Note that the derivative of $f(z)$ is given by $f'(z) = f(z)(1 - f(z))$.

The error back-propagation algorithm updates the weights w_i with $-c \frac{\partial J}{\partial w_i}$ (c is a constant). Show that the error correction algorithm of Q.1 is a simplified version of this algorithm by comparing the equations.

Q.4 Explain the reason why multi-layer networks that stack discriminant functions are effective for pattern classification. Show also that a two-layer network can be regarded as a weighted voting of results of multiple discriminant functions $(g_1(X), \dots, g_k(X))$.

ハッシュ法に関する以下の設問に答えよ。

設問1 ハッシュ法の原理を, 以下の用語をすべて用いて説明せよ。

ハッシュ表, ハッシュ関数, 衝突

設問2 衝突を解決する方法として, 開番地法的一种である線形走査法を用いる場合, データの探索アルゴリズム, 挿入アルゴリズム, 削除アルゴリズムをそれぞれ簡潔に記述せよ。

設問3 以下の6つの文字列をこの順に, サイズ10のハッシュ表に線形走査法によって格納することを考える。

tokyo tsu otsu kyoto osaka nara

ハッシュ関数を具体的に定義し, 設問2で解答した挿入アルゴリズムに従って各文字列がハッシュ表へ格納される過程を示せ。ただし, 各文字の内部コードは以下の表のと通りの整数値であるとする。

文字	内部コード	文字	内部コード	文字	内部コード	文字	内部コード
a	1	h	8	o	15	v	22
b	2	i	9	p	16	w	23
c	3	j	10	q	17	x	24
d	4	k	11	r	18	y	25
e	5	l	12	s	19	z	26
f	6	m	13	t	20		
g	7	n	14	u	21		

設問4 線形走査法ではクラスターという現象によって探索の効率が落ちる場合がある。この現象を説明し, さらに, これを回避する開番地法の別の方法を複数, 説明せよ。

Question is translated in English in the section below; this translation is given for reference only.

Answer the following questions about hashing.

Q.1 Explain the basic idea of hashing, using all following terms:

hash table, hash function, and collision.

Q.2 Suppose that we use linear probing, which is an open addressing method, to solve collision problems. Describe a lookup algorithm, an insertion algorithm, and a deletion algorithm for linear probing.

Q.3 Suppose that the following six strings are stored in this order in a hash table of size 10, using linear probing.

tokyo tsu otsu kyoto osaka nara

Define a hash function and show the process of storing these strings into a hash table, using the insertion algorithm that you described in the previous question. Use the character codes for the Roman alphabet given in the following table.

Character	Code	Character	Code	Character	Code	Character	Code
a	1	h	8	o	15	v	22
b	2	i	9	p	16	w	23
c	3	j	10	q	17	x	24
d	4	k	11	r	18	y	25
e	5	l	12	s	19	z	26
f	6	m	13	t	20		
g	7	n	14	u	21		

Q.4 The lookup performance of linear probing might go down due to a problem called clustering. Explain the clustering problem and other open addressing methods which can avoid it.

キャッシュメモリに関する以下の設問に答えよ。

設問 1 主記憶上のブロックをキャッシュ上に配置する方式として、直接マップ (direct mapped) 方式と完全連想 (fully associative) 方式がある。それぞれがどのような方式かを図を用いて説明せよ。

設問 2 完全連想方式においてキャッシュ・ミスが生じた際に置き換えるべきブロックを選択する手法 (policy) として、ランダム (random) 法、LRU (least-recently used) 法、FIFO (first-in first-out) 法がある。それぞれがどのような方式かを説明せよ。

設問 3 キャッシュのミス率は、主記憶の参照系列に依存する。完全連想方式において LRU 法が FIFO 法よりもミス率が小さくなる場合と大きくなる場合の参照系列の例をそれぞれ挙げよ。

設問 4 以下の仮定の下で、直接マップ方式と完全連想方式の CPU 時間を比較せよ。

- ・ CPI (1 実行命令あたりのクロック数)

直接マップ方式 1.5、完全連想方式 1.6

- ・ 1 実行命令あたりの平均メモリ参照回数 1.3

- ・ ミス率

直接マップ方式 0.04、完全連想方式 0.03

- ・ ミス・ペナルティ (キャッシュ・ミス時の主記憶参照に必要なクロック数) 10

ただし、キャッシュ・ヒット時にキャッシュの参照に必要な時間は CPI に含まれるものとする。

設問 5 小容量の命令専用キャッシュでは、直接マップ方式の方が完全連想方式で LRU 法を用いるよりも優れているといわれている。その理由について論ぜよ。

Question is translated in English in the section below; this translation is given for reference only.

Answer the following questions about CPU cache memories.

- Q.1 The fully-associative mapping and the direct mapping are algorithms for mapping memory blocks on the main memory onto the CPU cache. Explain each of these algorithms using figures.
- Q.2 The following policies are used when choosing a block to be replaced in a fully-associative CPU cache in case a cache miss occurs: the random policy, the LRU (least-recently used) policy, and the FIFO (first-in first-out) policy. Explain each of these policies.
- Q.3 The miss rate of a CPU cache memory depends on the sequence of memory access. Show an example of a sequence on which the miss rate in the LRU policy is less than that of the FIFO policy, and an example of a sequence on which the miss rate in the LRU policy is larger than that of the FIFO policy, in a fully-associative CPU cache.
- Q.4 Compare the CPU time with a direct-mapped cache and the CPU time with a fully-associative cache, under the following assumptions.
- CPI (Clock cycles Per Instruction):
Direct-mapped cache: 1.5, fully-associative cache: 1.6
 - Average number of memory access per instruction: 1.3
 - Miss rate:
Direct-mapped cache: 0.04, fully-associative cache: 0.03
 - Miss penalty (Number of required clock cycles in main memory access when a cache miss occurs): 10
- Assume that required clock cycles in cache access when a cache hit occurs are included in CPI.
- Q.5 It is said that the direct mapping is superior to the fully-associative mapping with the LRU policy for a small-size instruction cache. Discuss why.

整数 $n \geq 0$ 、実数 $-1 \leq x \leq 1$ に対して、

$$T_n(x) = \cos(n \cos^{-1} x) = \cos(n \arccos x)$$

としたとき、以下の設問に答えなさい。

設問 1 $T_n(x)$ は x の n 次多項式となり、 $n \geq 1$ に対して

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

が成り立つことを証明しなさい。

設問 2 $n \geq 1$ に対して $k = 1, 2, 3, \dots, n$ としたとき、 $T_n(x)$ の根 \hat{x}_k が

$$\hat{x}_k = \cos \frac{2k-1}{2n} \pi$$

となることを証明しなさい。

設問 3 実数 a, b ($\sin \frac{b}{2} \neq 0$) に対して

$$\sum_{m=0}^n \cos(a + mb) = \frac{\sin \frac{(n+1)b}{2} \cos \left(a + \frac{nb}{2} \right)}{\sin \frac{b}{2}}$$

となることおよび、設問 2 の等式を利用して、 $n \geq 1$ の時、整数 $0 \leq i, j < n$ に対して

$$\sum_{k=1}^n T_i(\hat{x}_k) T_j(\hat{x}_k) = \begin{cases} n & i = j = 0 \\ \frac{n}{2} & i = j \neq 0 \\ 0 & i \neq j \end{cases}$$

となることを証明しなさい。

Question is translated in English in the section below; this translation is given for reference only.

Define

$$T_n(x) = \cos(n \cos^{-1} x) = \cos(n \arccos x),$$

where n denotes an integer satisfying $n \geq 0$ and x a real variable satisfying $-1 \leq x \leq 1$. Answer the following questions.

Q.1 Prove that $T_n(x)$ can be represented by an n -th order polynomial of x and satisfies the following equation for $n \geq 1$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Q.2 Prove that for $n \geq 1$ and $k = 1, 2, 3, \dots, n$, all roots of $T_n(x)$ denoted by \hat{x}_k can be represented as

$$\hat{x}_k = \cos \frac{2k-1}{2n} \pi.$$

Q.3 For real numbers a and b such that $\sin \frac{b}{2} \neq 0$, the following equation holds

$$\sum_{m=0}^n \cos(a + mb) = \frac{\sin \frac{(n+1)b}{2} \cos \left(a + \frac{nb}{2} \right)}{\sin \frac{b}{2}}.$$

Using this equation and the equation in Q.2, prove that for $n \geq 1$ and integers i and j satisfying $0 \leq i, j < n$, the following equation holds

$$\sum_{k=1}^n T_i(\hat{x}_k) T_j(\hat{x}_k) = \begin{cases} n & i = j = 0 \\ \frac{n}{2} & i = j \neq 0 \\ 0 & i \neq j \end{cases}$$

n を正整数とするとき、整数 $1, 2, \dots, n$ を重複なく並べた列を置換列という。 $I = a_1, a_2, \dots, a_n$ を置換列とする。 $I' = a_{i_1}, a_{i_2}, \dots, a_{i_k}$ が $i_1 < i_2 < \dots < i_k$ を満たすとき、 I' は I の部分列であるという。 さらに $a_{i_1} < a_{i_2} < \dots < a_{i_k}$ を満たすとき、 I' を I の増加部分列、 $a_{i_1} > a_{i_2} > \dots > a_{i_k}$ を満たすとき、 I' を I の減少部分列という。 列 J に含まれる整数の個数を列 J の長さという（例えば、上記の I および I' の長さは、それぞれ n および k である）。 以下の設問に答えよ。

設問 1 $n = 10$ とし、 $I_1 = 4, 1, 10, 6, 7, 8, 2, 9, 3, 5$ を置換列とする。 I_1 の長さ 4 の増加部分列を 1 つ示せ。

設問 2 設問 1 の I_1 の最長増加部分列、最長減少部分列をそれぞれ 1 つずつ示せ。

設問 3 $1 \leq i \leq n$ の各 i について $s(i) = (c(i), d(i))$ を以下のように定義する。

$c(i)$: a_i で始まる I の増加部分列のうち、最長のものの長さ

$d(i)$: a_i で始まる I の減少部分列のうち、最長のものの長さ

設問 1 の I_1 について、 $s(4)$, $s(7)$, $s(9)$ を求めよ。

設問 4 任意の置換列について、 $i \neq j$ ならば $s(i) \neq s(j)$ となることを示せ。

設問 5 正整数 t に対して $n = t^2 + 1$ とするとき、長さ n の任意の置換列は長さ $t + 1$ の増加部分列または長さ $t + 1$ の減少部分列を持つことを示せ。

設問 6 長さ 4 の増加部分列も長さ 4 の減少部分列も持たない、長さ 9 の置換列を 1 つ示せ。

Question is translated in English in the section below; this translation is given for reference only.

For a positive integer n , a sequence obtained by arranging integers $1, 2, \dots, n$ without repetition is called a permutation sequence. Let $I = a_1, a_2, \dots, a_n$ be a permutation sequence. If $I' = a_{i_1}, a_{i_2}, \dots, a_{i_k}$ satisfies $i_1 < i_2 < \dots < i_k$ then I' is called a subsequence of I . If further $a_{i_1} < a_{i_2} < \dots < a_{i_k}$ holds, then I' is called an increasing subsequence of I , and if $a_{i_1} > a_{i_2} > \dots > a_{i_k}$ holds, then I' is called a decreasing subsequence of I . The length of a sequence J is the number of integers contained in J (for example, the lengths of the above I and I' are n and k , respectively). Answer the following questions.

- Q.1 Let $n = 10$ and $I_1 = 4, 1, 10, 6, 7, 8, 2, 9, 3, 5$ be a permutation sequence. Give an increasing subsequence of I_1 of length four.
- Q.2 Give a longest increasing subsequence and a longest decreasing subsequence of I_1 in Q.1.
- Q.3 For each i ($1 \leq i \leq n$), we define $s(i) = (c(i), d(i))$, where $c(i)$ is the length of the longest increasing subsequence(s) of I that start(s) with a_i , and $d(i)$ is the length of the longest decreasing subsequence(s) of I that start(s) with a_i . Compute $s(4)$, $s(7)$, and $s(9)$ for I_1 in Q.1.
- Q.4 Show that, for every permutation sequence, $s(i) \neq s(j)$ holds if $i \neq j$.
- Q.5 Let $n = t^2 + 1$ for a positive integer t . Show that every permutation sequence of length n has an increasing subsequence of length $t + 1$ or a decreasing subsequence of length $t + 1$.
- Q.6 Give a permutation sequence of length nine that has neither an increasing subsequence of length four nor a decreasing subsequence of length four.

設問1 壺の中に赤玉が u 個, 白玉が v 個入っている. この壺から n 個の玉を一度に無作為に取り出すとき, 以下の問いに答えよ. ただし, X は取り出した玉の中に入っている赤玉の個数を表す確率変数とする.

(1) 取り出した玉の中に k 個の赤玉が入っている確率 $P(X = k)$ を求めよ.

(2) k のとりうる最小値 s と最大値 t を求めよ.

(3) (1) で求めた確率が $\sum_{k=s}^t P(X = k) = 1$ を満たすことを計算により示せ.

設問2 池の中の魚の総数 N を推定するために, u 匹の魚を捕獲してマークを付けて池に戻す. しばらくたってから, n 匹の魚を捕獲したところ, その中の k 匹にマークがついていた. このとき, 池の中にいる魚の総数 N の最尤推定値を計算せよ.

Question is translated in English in the section below; this translation is given for reference only.

Q.1 There are u red balls and v white balls in a pot. n balls are randomly drawn at once from the pot. X is a random variable that counts the number of red balls in the random draw.

- (1) What is the probability $P(X = k)$ that k balls are red in a draw?
- (2) Find the minimum value s and maximum value t of k .
- (3) Prove that $\sum_{k=s}^t P(X = k) = 1$ holds using your answer to (1) and mathematical deduction.

Q.2 In order to estimate the total number of fish N in a pond, we catch u fish, mark them and release them in the pond. Later, we catch n fish from the pond and find k marked fish among them. Calculate the maximum likelihood estimate for N .