

平成 26 年度 10 月期入学 / 平成 27 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 26 年 8 月 6 日 13:00~16:00

【注意】

- 問題冊子はこの表紙を含めて 33 枚ある。
- 試験開始の合図があるまで中を見てはいけない。
- 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 |

- 特に指定のない限り、日本語または英語で解答すること。
- 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 2014 Admissions / April 2015 Admissions
Entrance Examination for Master's Program
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
(Specialized Subjects)

August 6, 2014
13:00 - 16:00

NOTES

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

| | |
|--|---------------|
| Bioinformatics (Question Number B-1 to B-4) | Page 1 to 8 |
| Psychology (Question Number P-1 to P-4) | Page 9 to 16 |
| Computer science, Electrical and electronic engineering (Question Number T-1 to T-8) | Page 17 to 32 |

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

野生型のマウスでは音刺激（パルス刺激）による驚愕反応は、先行する微弱な音刺激（プレパルス刺激）によって抑制される。このプレパルスによるパルス刺激に対する驚愕反応の抑制をプレパルス抑制と呼ぶ。遺伝子 x にナンセンス変異をもつマウスの変異体を樹立し、聴覚性驚愕反応を解析した。図は野生型と変異体におけるパルス刺激による驚愕反応の測定結果（図1）、プレパルス抑制の測定結果（図2）を示す。遺伝子 x は、一部のグルタミン酸作動性神経細胞から投射を受ける神経細胞に発現していた。また、変異体において遺伝子 x を発現している細胞は、正常な電気刺激による神経活動を示した。

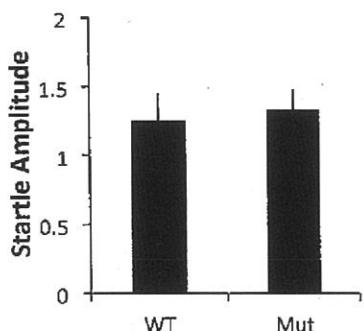


図1: パルス刺激による驚愕反応
WT:野生型 Mut:変異体

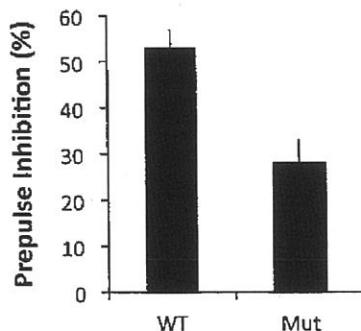


図2: 驚愕反応のプレパルス抑制
WT:野生型 Mut:変異体

- 設問1 驚愕反応の測定方法について述べよ。
設問2 プレパルス抑制の認知機能との関連性について述べよ。
設問3 ナンセンス変異について説明せよ。
設問4 聴覚性驚愕反応の神経回路を、プレパルス刺激はどのように制御していると考えられるか述べよ。
設問5 遺伝子 x の遺伝子 x 発現細胞における役割について述べよ。

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

Weak acoustic prestimuli (prepulse) can inhibit the startle response, which is usually induced by acoustic stimuli (pulse), in wildtype mice. This attenuation of acoustic startle response by prepulse stimuli is called a prepulse inhibition. Mutant mice, whose gene x has undergone a nonsense mutation, were grown and their acoustic startle response was subsequently tested. Figure 1 shows acoustic startle response by pulse stimuli. Figure 2 shows prepulse inhibition. The gene x is expressed in a part of neurons, which are projected from glutamatergic neurons. The gene x -expressing cells in mutant mice showed normal neural response to electrical stimuli.

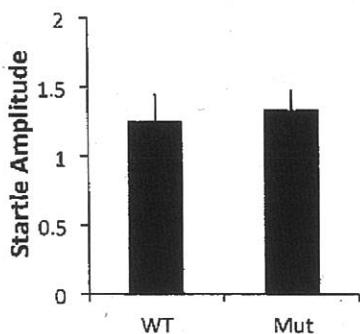


Figure 1. Startle response by pulse stimuli. WT: Wildtype mice. Mut: Mutant mice.

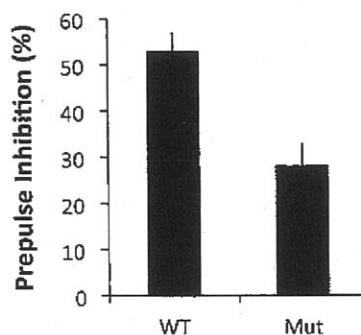


Figure 2. Prepulse Inhibition. WT: Wildtype mice. Mut: Mutant mice.

- Q.1 Describe the measuring method of acoustic startle response.
- Q.2 Describe the relevance of prepulse inhibition to cognitive function.
- Q.3 Explain what "nonsense mutation" means.
- Q.4 Describe a role of prepulse stimuli on the neural circuit of acoustic startle response.
- Q.5 Describe the role of gene x in gene x -expressing cells.

ある真核細胞において、遺伝子 x, y の転写を解析した。遺伝子 x, y がコードするタンパク質 X, Y は、それぞれ遺伝子 y, x の転写を抑制していた（図 1, 2）。

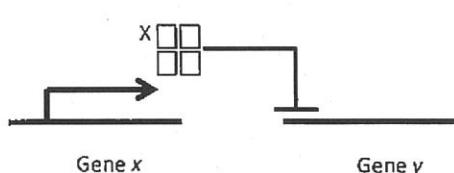


図 1: 遺伝子 x にコードされたタンパク質 X は遺伝子 y の転写を抑制する。

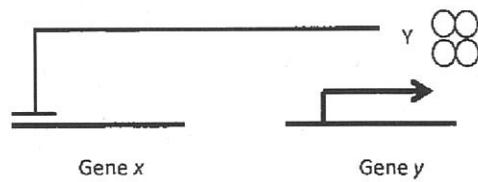


図 2: 遺伝子 y にコードされたタンパク質 Y は遺伝子 x の転写を抑制する。

設問 1 遺伝子の転写を活性化する転写因子は、どのようにして遺伝子の転写を活性化しているか。簡単に説明せよ。

設問 2 遺伝子の転写を測定する方法をひとつあげ、説明せよ。

設問 3 遺伝子 y が発現している状態の細胞が細胞分裂をおこなった。分裂した細胞では遺伝子 x および遺伝子 y はどのような状態にあるか、理由とともに述べよ。

設問 4 転写因子を介さずに転写が制御されるしくみをひとつあげ、説明せよ。

設問 5 細胞に遺伝子 x の発現を誘導した。遺伝子 x 、遺伝子 y の遺伝子発現の時間的挙動について述べよ。

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

Transcription of gene x and gene y was analyzed in a eucaryocyte. Protein X and protein Y, which are encoded by gene x and gene y , repress transcription of gene y and gene x , respectively (Figures 1 and 2).

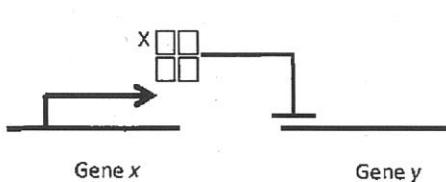


Figure 1. Protein X encoded by gene x represses gene y transcription.

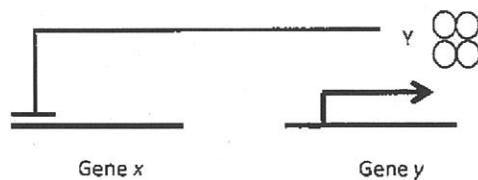


Figure 2. Protein Y encoded by gene y represses gene x transcription.

- Q.1 Describe briefly how transcription factors, which activate gene transcription, activate transcription.
- Q.2 Describe in detail a measuring method for gene transcription.
- Q.3 Suppose that a cell expressing gene y has been divided. Describe the status of gene x and gene y in the daughter cells.
- Q.4 Describe in detail a mechanism to control gene expression without using transcription factors.
- Q.5 Describe the temporal behavior of gene x expression and gene y expression when gene x is induced in a cell.

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てもよい。

1. 制限酵素断片長多型 (RFLP)
2. クリスパー (CRISPR)
3. ホメオボックス
4. イントロン
5. PDB (Protein Data Bank)
6. PROSITE データベース
7. BLAST
8. ホモロジー検索
9. 最大節約法
10. Hill 式

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

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

1. Restriction Fragment Length Polymorphism (RFLP)
2. Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)
3. Homeobox
4. Intron
5. PDB (Protein Data Bank)
6. PROSITE Database
7. BLAST
8. Homology Search
9. Maximum Parsimony Method
10. Hill Equation

DNA配列のパターンを表現するために、長さ0以上の任意のDNA配列とマッチする文字^(*)とA,C,G,Tの合計で5種類の文字からなるパターン文字列を考える。ここで、 $F(s, n)$ をパターン文字列 s とマッチする長さ n のDNA配列すべての集合とする。例えば、 $s = A^*CGT$ はACGT, AACGT, ACACGTのいずれともマッチし、 $s = A^*G*T$ はAGT, ACGT, AGCT, ACGCTのいずれともマッチする。また、

$$\begin{aligned} F(A^*CGT, 5) &= \{AACGT, ACCGT, AGCGT, ATCGT\} \\ F(A^*G^*, 3) &= \{AAG, ACG, AGG, ATG, AGA, AGC, AGT\} \end{aligned}$$

が成立し、 $|F(A^*CGT, 5)| = 4$, $|F(A^*G^*, 3)| = 7$ となる（ $|S|$ は集合 S の要素数を表す）。以下の設問に答えよ。

設問1 $|F(*A^*T, 3)|$ を求めよ。

設問2 $|F(A^*G^*T, 5)|$ を求めよ。

設問3 $|F(A^*T, n)|$ を与える n に関する簡潔な式を理由とともに示せ（ただし、 $n \geq 2$ ）。

設問4 下記は $|F(A^*C^*G^*T, n)|$ を与える n に関する式である（ただし、 $n \geq 4$ ）。長方形の空欄2箇所を埋めてこの式を完成させ、理由とともに示せ。

$$\sum_{i=3}^{\square} \square$$

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

In order to represent patterns in DNA sequences, we consider pattern strings over an alphabet $\{A,C,G,T,*\}$, where '*' is a special letter that matches any DNA sequence (i.e., any string over $\{A,C,G,T\}$) including the empty string. Let $F(s, n)$ be the set of DNA sequences of length n that match a pattern string s . For example, $s = A^*CGT$ can match any of ACGT, AACGT, and ACACGT, and $s = A^*G*T$ can match any of AGT, ACGT, AGCT, and ACGCT. Furthermore, the following equalities hold:

$$\begin{aligned}F(A^*CGT, 5) &= \{AACGT, ACCGT, AGCGT, ATCGT\}, \\F(A^*G^*, 3) &= \{AAG, ACG, AGG, ATG, AGA, AGC, AGT\},\end{aligned}$$

and thus we have $|F(A^*CGT, 5)| = 4$ and $|F(A^*G^*, 3)| = 7$, where $|S|$ denotes the number of elements in a set S .

Answer the following questions.

Q.1 List all elements of $F(*A^*T, 3)$.

Q.2 Compute $|F(A^*G^*T, 5)|$.

Q.3 Describe a simple mathematical formula on n that gives $|F(A^*T, n)|$ and explain its derivation, where $n \geq 2$.

Q.4 The following is a mathematical formula on n that gives $|F(A^*C^*G^*T, n)|$, where $n \geq 4$. Complete the formula by filling the blank boxes, and explain its derivation.

$$\sum_{i=3}^{\square} \boxed{\quad}$$

以下の設間に答えよ。

Non-Disclosure

[Liepelt, R., and Brass, M. (2010) Top-down modulation of motor priming by belief about animacy. *Experimental Psychology*, 57, 221-227.を一部改変]

- 設問 1 本実験で “group” と “congruency”とは何か説明せよ。
- 設問 2 本実験で “motor priming”とは何か。具体的に述べよ。
- 設問 3 下線 ① の意味を具体的に述べよ。
- 設問 4 下線 ② の結果をミラーニューロンシステムもしくは行為の実行/観察システムとの関係において考察せよ。

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

Answer the following questions.

Non-Disclosure

[Modified from Liepelt, R., and Brass, M. (2010) Top-down modulation of motor priming by belief about animacy. *Experimental Psychology*, 57, 221-227.]

- Q.1 What are the “group” and the “congruency” in this experiment?
- Q.2 What is the “motor priming” in this experiment? Give an explanation in detail.
- Q.3 What does the underlined part ① mean? Give an explanation in detail.
- Q.4 What does the underlined sentence ② suggest? Explain it from the viewpoint of mirror neuron system or action execution/observation matching system.

以下の設問に答えよ。

Non-Disclosure

[Krueger, L. (1984) Perceived numerosity: A comparison of magnitude production, magnitude estimation, and discrimination judgments. *Perception & Psychophysics*, 35, 536-542.を一部改変]

- 設問 1 power-function exponent とは何かを述べよ。
- 設問 2 subjective magnitude of numerosity はどのようにして得られるか具体的に述べよ。
- 設問 3 numerosity におけるjndsを求めたい。どのようにすればよいか述べよ。

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

Answer the following questions.

Non-Disclosure

[Modified from Krueger, L. (1984) Perceived numerosity: A comparison of magnitude production, magnitude estimation, and discrimination judgments. *Perception & Psychophysics*, 35, 536-542.]

- Q.1 Explain what the power-function exponent is.
- Q.2 How is the subjective magnitude of numerosity obtained?
- Q.3 How are the jnds of numerosity obtained? Give the procedure in detail.

Figure 1 は「注意の瞬き」と呼ばれる現象を調べるための実験の刺激系列を図示したものである。英文字あるいは数字が 1 秒間に 10 文字の速さで提示される。実験参加者の課題は、文字列の中から、第 1 ターゲット (T1; 緑色の文字) を報告し、さらに第 2 ターゲット (T2; 数字) を報告することであった (T1 and T2 report condition)。また、統制条件として、T1 and T2 report condition と同じ刺激系列を用い、第 2 ターゲットのみを報告する実験ブロック (T2 report condition) が設けられた。Figure 2 は、この実験の結果を示したものである。第 2 ターゲットの正答率が、T1 and T2 lag の関数としてプロットされている。T1 and T2 lag とは、第 1 ターゲットの後、何文字目に第 2 ターゲットが表示されたかを示す。

以下の設問に答えよ。

設問 1 「注意の瞬き」が、第 2 ターゲットに対する視覚マスキングでは解釈できない理由を説明せよ。

設問 2 「注意の瞬き」のメカニズムを、二重課題と心的資源の用語を用いて説明せよ。

設問 3 「注意の瞬き」のメカニズムを、ワーキングメモリと記憶の固定化の用語を用いて説明せよ。

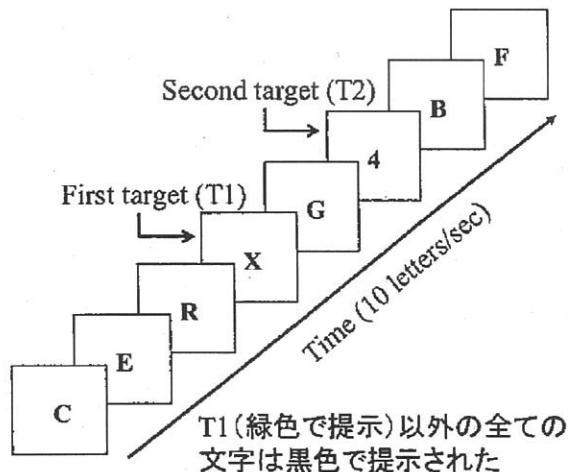


Figure 1

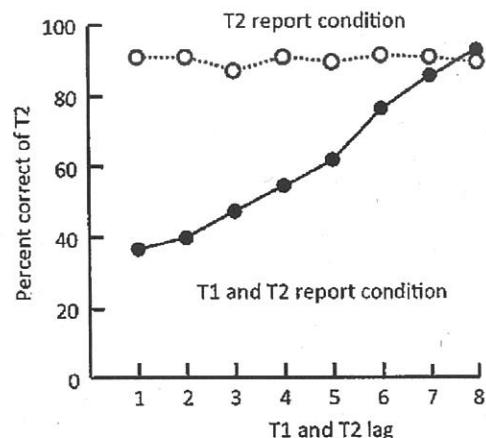


Figure 2

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

Figure 1 shows a sequence of stimuli used in an experiment for examining a phenomenon referred to as "attentional blink". Alphabetic or numeric letters were presented in a rate of 10 letters/sec. In the experiment, a participant was asked to report the first target (T1, an alphabetic letter presented in green) and then report the second target (T2, a numeric letter). This is named the T1 and T2 report condition. With the same stimulus sequence, an experimental block requiring only the report of T2 is named the T2 report condition. Figure 2 shows a result of the experiment, in which the accuracy of T2 report is plotted as a function of T1 and T2 lag (T1 and T2 lag refers to the temporal difference of T1 and T2 counted by the number of appeared letters).

Answer the following questions.

- Q.1 Explain the reason why attentional blink cannot be interpreted in terms of the visual masking to T2.
- Q.2 Explain the mechanism of attentional blink using following terms: "dual-task" and "mental resource".
- Q.3 Explain the mechanism of attentional blink using following terms: "working memory" and "memory consolidation".

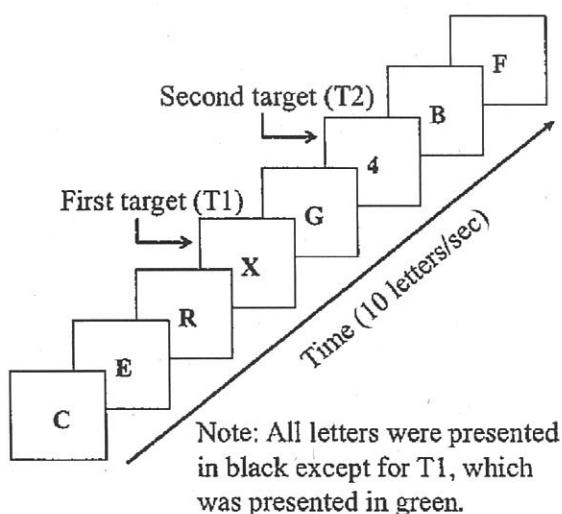


Figure 1

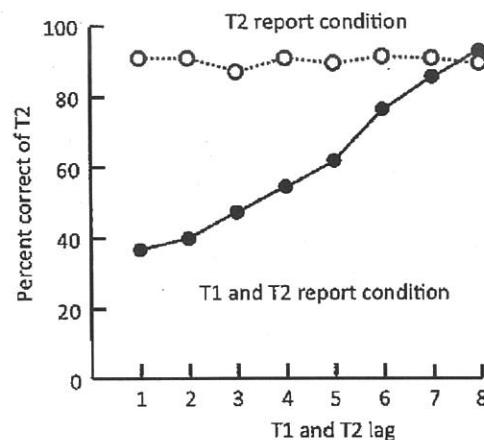


Figure 2

Figure 1 は、ある脳機能計測 (fMRI) 研究におけるストループ課題の結果を示したものである。この課題では、ある色で描かれた色名単語がターゲットとして提示された（例えば、緑色で描かれた red という単語）。実験参加者の課題は、ターゲットの色を答えることであった。Figure 1 には、直前の試行のタイプの関数として、現試行のタイプ別の反応時間 (RT) がプロットされている。trial type C はターゲットの色名と文字色が一致している場合、trial type I はターゲットの色名と文字色が不一致の場合を指す。

以下の設問に答えよ。

設問 1 current trial type C と current trial type I の反応時間の差は一般にストループ効果として知られている。ストループ効果がなぜ起るのかについて説明せよ。

設問 2 Figure 1 のデータを検定したところ、previous trial type と current trial type の交互作用が有意であった。この交互作用が意味することを説明せよ。

設問 3 Figure 2 は、current trial type I のときの脳の前部帯状皮質 (ACC) の活動 (BOLD 信号) の変化率 (% signal change) を、current trial type C のときをベースラインとして算出したものである。previous trial type C の場合 (cI trial sequence) と I の場合 (iI trial sequence) に分けてプロットされている。1 秒のところでターゲットが提示された。Figure 2 からわかる ACC の働きに基づいて、Figure 1 の結果を説明せよ。

Non-Disclosure

(Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303, 1023–1026. を改変)

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

Figure 1 shows a result of a Stroop task, obtained in a brain imaging (fMRI) study. In the task, a colored color word was presented as a target (e.g., a word "red" appeared green). A participant's task was to report the appearing color of the target. In Figure 1, mean reaction times (RTs) to targets were plotted as a function of the previous trial type, separately for each current trial type. Trial type *C* refers to the trial in which the name of a color word is congruent with its appearing color, and trial type *I* refers to the trial in which the name of a color word is incongruent with its appearing color.

Answer the following questions.

Q.1 The difference in RTs between the current *C* trial and the current *I* trial is known as the Stroop effect. Explain why the Stroop effect occurs.

Q.2 There was a significant interaction between the previous trial type and the current trial type in the data shown in Figure 1. Explain what the interaction implies.

Q.3 Figure 2 shows a % signal change of BOLD signal (the response in the anterior cingulate cortex (ACC) in the brain) in the current *I* trial comparing with the current *C* trial as a baseline. The % signal change in the previous *C* trial (cI trial sequence) and that in the previous *I* trial (iI trial sequence) are depicted separately. A target was presented at the time point of 1 sec of the figure. Explain the result of Figure 1 in terms of the function of ACC found in Figure 2.

Non-Disclosure

(Adapted from Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303, 1023–1026.)

$M \geq 2$ なる M 個の情報源記号 A_i ($i = 1, \dots, M$) をもつ無記憶情報源 S が与えられたとする。 S の各情報源記号 A_i に対して次のように 2 元符号語 C_i を定める。

- (1) 情報源記号 A_i の生起確率を p_i (ただし, $p_i > 0$) とするとき, α_i ($i = 1, \dots, M$) を次のように定める。

$$\alpha_1 = \frac{1}{2} p_1, \quad \alpha_i = \frac{1}{2} p_i + \sum_{k=1}^{i-1} p_k \quad (2 \leq i \leq M)$$

- (2) 不等式

$$2^{1-\alpha_i} \leq p_i < 2^{2-\alpha_i}$$

を満足する整数 l_i に対して, α_i を 2 進小数展開したときの小数点以下 l_i 術の 2 進数表示を C_i とする。

このとき, 次の問いに答えなさい。ここで符号アルファベットは $\{0, 1\}$ とする。

設問 1 : 生起確率が $0.18, 0.08, 0.02, 0.04$ の情報源記号に対して, 上記の符号化の方式で得られる符号語を示しなさい。

設問 2 : 上記の符号化の方式で得られる符号が瞬時符号であることを示しなさい。

設問 3 : 上記の符号化の方式で得られる符号の 1 情報源記号あたり平均符号長の範囲を S のエントロピー $H(S)$ を用いた不等式として示しなさい。

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

Let M be an integer no less than 2, and consider a memoryless information source S that has alphabet symbols A_i ($i = 1, \dots, M$) and a procedure of associating binary code word C_i for A_i as follows:

- (1) Let $p_i > 0$ be the probability of occurrence of symbol A_i , and α_i ($i = 1, \dots, M$) be:

$$\alpha_1 = \frac{1}{2} p_1, \quad \alpha_i = \frac{1}{2} p_i + \sum_{k=1}^{i-1} p_k \quad (2 \leq i \leq M).$$

- (2) The code word C_i for A_i is given as the binary sequence of length l_i after the binary point obtained from binary expansion of α_i to the l_i -th place, where l_i is an integer that satisfies $2^{1-l_i} \leq p_i < 2^{2-l_i}$.

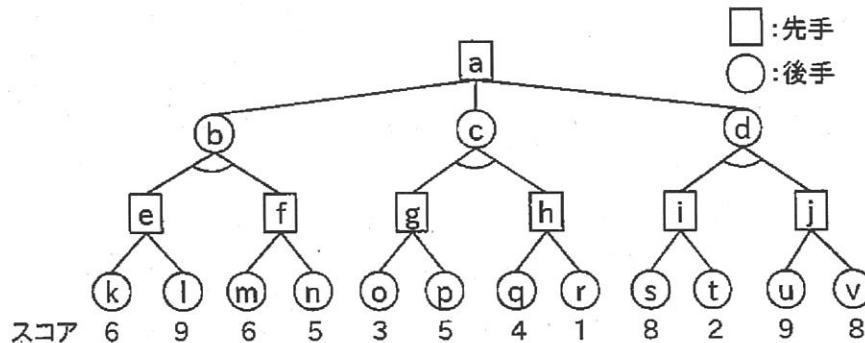
Let the code alphabet be $\{0, 1\}$ and answer the following questions.

Q.1 What are the code words that the above procedure specifies for source symbols whose probabilities of occurrence are 0.18, 0.08, 0.02, and 0.04?

Q.2 Show that the above procedure will yield an instantaneous code.

Q.3 Display the range of average length of code words per source symbol as one or more inequalities using the entropy of S , $H(S)$.

二人のプレイヤーが交互に手を指すゲームを考える。下図は、このゲームで現局面（ノード a）から三手先までの局面を表したゲーム木である。図の葉ノード下の数字は、三手先の局面での先手の有利さを表すスコアである。三手先の局面のスコアをもとに先手の次の手を決める。この時、以下の設間に答えよ。



設問1 minimax 法で先手の最善手を選ぶ時、ノード a~j の評価値 (minimax 値) を示せ。また、この時に選択される先手の次の手を答えよ。

設問2 上図の左の子ノードから順に α - β 法で枝刈りしながらゲーム木を探索する。この時、 α カット、 β カットにより枝刈りされるノードをそれぞれ全て答えよ。

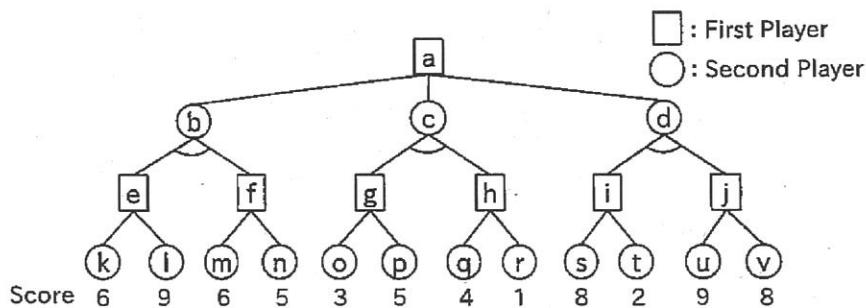
設問3 一般に、 α - β 法で枝刈りしながらゲーム木を探索する時、最も多数のノードを枝刈りできるのはどのような場合か、説明せよ。

設問4 上図をある順序で探索した時、 α - β 法で最も多数のノードを枝刈りできたとする。この時、 α カット、 β カットにより枝刈りされるノードをそれぞれ全て答えよ。

設問5 一般に、 α - β 法で最も多数のノードを枝刈りできるゲーム木の探索順序は、事前には分からない。この時、どのような方法で探索順序を決定すれば、 α - β 法で、より多数の枝刈りが行えると期待できるか、説明せよ。

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

The graph below shows the partial game tree of the game from the current state (node a) to all possible states after 3 moves. The number below each leaf node denotes the score of advantage earned by the first player after 3 moves. The first player decides his/her first move using the information displayed in the tree below. Answer the questions below:



Q.1 Describe the valuation of each node a, ..., j if the first player uses the minimax algorithm to decide his/her moves. What is the best move of the first player at the current state in the minimax sense?

Q.2 Suppose the game tree above is traversed from left to right, and that it is pruned using the α - β algorithm. List the nodes pruned by α -cut and those pruned by β -cut respectively in this case.

Q.3 In which situations is the α - β algorithm efficient, in the sense that it can prune out a maximal number of nodes?

Q.4 Consider the traversal order which maximizes, for the above tree, the number of pruned nodes using the α - β algorithm. List the nodes pruned by α -cut and those pruned by β -cut respectively when traversing the tree with that traversal order.

Q.5 For a given game tree, it is difficult, in general, to find the traversal order which maximizes the number of pruned nodes using the α - β algorithm. Discuss a method to choose a traversal order that is expected to prune a larger number of nodes.

設問 1 2 クラスの 2 次元データ群 A, B をクラス識別する手法について考える。ただし、A, B の教師（訓練）データはそれぞれ $\left\{ \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \dots, \begin{bmatrix} a_{1M} \\ a_{2M} \end{bmatrix} \right\}$, $\left\{ \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}, \dots, \begin{bmatrix} b_{1M} \\ b_{2M} \end{bmatrix} \right\}$, 平均はそれぞれ $\bar{\mathbf{m}}_A = \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \end{bmatrix}$, $\bar{\mathbf{m}}_B = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \end{bmatrix}$ であるとする。

(1) A の分散共分散行列 Σ_A を求めよ。

(2) A の相関行列 $\Sigma'_A = \frac{1}{M} \begin{bmatrix} \sum_{k=1}^M (a_{1k})^2 & \sum_{k=1}^M a_{1k}a_{2k} \\ \sum_{k=1}^M a_{1k}a_{2k} & \sum_{k=1}^M (a_{2k})^2 \end{bmatrix}$ を Σ_A および \mathbf{m}_A で表せ。

設問 2 設問 1において $\mathbf{m}_A = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$, $\mathbf{m}_B = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\Sigma_A = \Sigma_B = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$ であるとする。
このとき、入力点 $x_1 = \begin{bmatrix} -4 \\ 0 \end{bmatrix}$, $x_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$, $x_3 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ がどちらのクラスに属するか求めたい。

- (1) フィッシャーの線形判別により、 x がクラス A に属するときに $f(x) > 0$, クラス B に属するときに $f(x) < 0$ となる線形識別関数 $f(x) = w_0 + \mathbf{w}^T x$ を求め、入力点 x_1 , x_2 , x_3 の識別を行え。計算の過程も示すこと。
- (2) 部分空間法の一種である CLAFIC (class-featuring information compression) 法を用いて入力点 x_1 , x_2 , x_3 の識別を行え。ただし CLAFIC 法とは相関行列の固有ベクトルと入力点との射影の大きさを用いて識別を行う方法である。ここでは第 1 固有ベクトルのみを用いる事とし、計算の過程も示すこと。ただし、 M が十分大きいと見なし $\frac{M-1}{M} \approx 1$ を用いてよい。
- (3) 設問 2 (1)(2) の結果を踏まえ、フィッシャーの線形判別と部分空間法 (CLAFIC 法) の違いについて論ぜよ。

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

Q.1 Consider the problem of obtaining a binary linear classifier that classifies two-dimensional data into classes **A** and **B**. Let the training data of classes **A** and **B** be $\left\{ \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \dots, \begin{bmatrix} a_{1M} \\ a_{2M} \end{bmatrix} \right\}$ and $\left\{ \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}, \dots, \begin{bmatrix} b_{1M} \\ b_{2M} \end{bmatrix} \right\}$, and their means be $\mathbf{m}_A = \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \end{bmatrix}$ and $\mathbf{m}_B = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \end{bmatrix}$.

- (1) Represent the covariance matrix Σ_A of class **A**.
- (2) Represent the correlation matrix $\Sigma'_A = \frac{1}{M} \begin{bmatrix} \sum_{k=1}^M (a_{1k})^2 & \sum_{k=1}^M a_{1k}a_{2k} \\ \sum_{k=1}^M a_{1k}a_{2k} & \sum_{k=1}^M (a_{2k})^2 \end{bmatrix}$ of class **A** using Σ_A and \mathbf{m}_A .

Q.2 In Q.1, let $\mathbf{m}_A = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$, $\mathbf{m}_B = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and $\Sigma_A = \Sigma_B = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$. Now we want to classify the points $\mathbf{x}_1 = \begin{bmatrix} -4 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

- (1) In accordance to the Fisher's linear discriminant, obtain the linear discriminant function $f(\mathbf{x}) = w_0 + \mathbf{w}^T \mathbf{x}$, which satisfies $f(\mathbf{x}) > 0$ when \mathbf{x} is in the class **A** and $f(\mathbf{x}) < 0$ when \mathbf{x} is in the class **B**, then classify the points \mathbf{x}_1 , \mathbf{x}_2 , and \mathbf{x}_3 . Give the procedure in detail.
- (2) Classify the points \mathbf{x}_1 , \mathbf{x}_2 , and \mathbf{x}_3 using the subspace classifier CLAFIC (class-featuring information compression) method, where the CLAFIC method uses the length of the projection between the data point and the eigenvector(s) of the correlation matrix. Here, use only the first eigenvector and assume $M \gg 1$, therefore $\frac{M-1}{M} \approx 1$. Give the procedure in detail.
- (3) Discuss the differences of the Fisher's linear discriminant and the CLAFIC method considering the results of Q.2(1) and Q.2(2).

設問1 10以上、99以下の整数が n 個入った配列 A について、配列の要素を昇順に整列することを考える。このとき、次の問いに答えよ。

- 配列 A が[34, 70, 13, 55, 21, 89, 41]であるときに、これに対するヒープを図示せよ。
- (a) の配列 A をヒープソートによって整列する手順について、時間複雑度がどのようになるかを含めて説明せよ。
- ある配列 A を整列するアルゴリズムのうち、時間複雑度が $O(n)$ のものを一つ挙げて、空間複雑度がどのようになるかを含めて、そのアルゴリズムを説明せよ。

設問2 整列済みの整数を要素とする配列において、ある整数が含まれているかどうかを調べる効率的なアルゴリズムを一つ挙げて、時間複雑度がどのようになるかを含めて説明せよ。

設問3 英単語のリストがあるときに、これからアナグラムの関係にある単語集合をすべて列挙する効率的な方法を考えて説明せよ。ただし、アナグラムとは、たとえば post と spot の関係のように、ある単語を構成する文字を異なる順番で使った単語を意味する。

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

Q.1 Consider an array A that contains n integers ranging from 10 to 99. Suppose that we sort the elements of A in ascending order. Answer the following questions.

- (a) When A is [34, 70, 13, 55, 21, 89, 41], draw a heap for this array.
- (b) Explain the procedure of heap sort for A defined in (a) and the time complexity of this algorithm.
- (c) Describe an $O(n)$ -time algorithm for sorting an array A . Explain the space complexity of the algorithm.

Q.2 Describe an efficient algorithm for finding an integer from an array that contains sorted integers. Explain the time complexity of the algorithm.

Q.3 When there is a list of English words, describe an efficient method for listing all sets of anagrams. An anagram means a word that is formed by changing the order of the letters of another word. For example, "post" is an anagram of "spot."

キャッシュメモリに関する以下の設問に答えよ。

設問1 主記憶とキャッシュメモリのマッピングには、完全連想方式、直接マップ方式、群連想方式があり、完全連想方式と直接マップ方式は群連想方式の特別な場合と見なすことができる。群連想方式がどのような方式か、図を用いて説明せよ。また完全連想方式と直接マップ方式それぞれについて、群連想方式のどのような場合に相当するか、またその長所・短所を説明せよ。

設問2 キャッシュメモリが 128K バイト、実アドレスが 32 ビットの計算機で、ブロックの大きさが 32 バイトの 8 ウェイ群連想方式を採用したとき、群番号及びタグ（群内ブロック番号）の指定に何ビット必要か説明せよ。

設問3 キャッシュメモリからの読み出しが、単純にはタグを用いてディレクトリを検索した後に、一致したエントリに対応するブロック枠にアクセスすることで実現される。これをより高速に行う手法の 1 つである、ディレクトリの検索とブロックからの読み出しを並列に行う手法について、設問 2 の 8 ウェイ群連想方式を例に説明せよ。

ページ方式仮想記憶に関する以下の設問に答えよ。

設問4 ページ方式仮想記憶における仮想アドレスから実アドレスへのアドレス変換にはページテーブルが用いられる。具体的なアルゴリズムについて説明するとともに、なぜキャッシュメモリで用いられた方が仮想記憶のアドレス変換に用いられないのか説明せよ。

設問5 多段ページテーブル方式の概要とその利点を、図を用いて説明せよ。

設問6 ページ方式仮想記憶とキャッシュメモリを併用した場合、仮想アドレスを実アドレスに変換してキャッシュメモリにアクセスする方式と、逆に仮想アドレスのままキャッシュメモリにアクセスする方式が考えられる。後者はアドレス変換無しにキャッシュにアクセスできるが、性能低下要因もある。どのような性能低下要因が考えられるか述べよ。

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

Answer the following questions about CPU cache memory.

- Q.1 The fully-associative mapping, the direct mapping, and the set associative mapping are algorithms mapping between the main memory and the CPU cache. The fully-associative mapping and the direct mapping can be regarded as special cases of the set associative mapping. Explain the set associative mapping algorithm with figures, and explain how to realize the fully-associative and the direct mappings by specializing the set associative mapping. Describe their respective advantages and disadvantages.
- Q.2 Consider a system of 128 KB cache and 32-bit physical address utilizing an 8-way set associative mapping with 32-byte block size. How many bits are required for the set number and the tag (block number inside a set)?
- Q.3 The cache read process can be realized by searching the directory to find an entry having the tag of interest first, and then by accessing the cache block corresponding to the entry. Explain how to speed up this process by parallelizing the search and the block access, by using the 8-way set associative example of Q.2.

Answer the following questions about paged virtual memory.

- Q.4 Virtual-to-physical address translations in paged virtual memory systems can be realized using a page table. Explain the algorithm, and explain why the mapping algorithms used for the CPU cache are not suitable for virtual-to-physical address translations.
- Q.5 Explain the multilevel page table architecture and its advantages using figures.
- Q.6 When used in combination with CPU cache memory, the system can translate virtual addresses into physical addresses to access the cache, or the system can use the virtual addresses as-is. The latter approach can omit the virtual-to-physical address translation, but has other factors making the system less efficient. Describe a factor that can spoil the system performance in the latter approach.

複素数 $z = x + iy$ の関数 $f(z) = z^2 + 2z$ について以下の問いに答えよ。ただし、 x, y は実数であり、 i は虚数単位である。

設問 1 $f(z)$ の実部 $u(x, y)$ と虚部 $v(x, y)$ を求めよ。

設問 2 $u(x, y)$ と $v(x, y)$ が次のコーシー・リーマンの微分方程式を満たすことを示せ。

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$$

設問 3 複素数 $\bar{z} = x - iy$ の関数 $g(\bar{z}) = \bar{z}^2 + 2\bar{z}$ がコーシー・リーマンの微分方程式を満たさないことを示せ。

設問 4 コーシー・リーマンの微分方程式を満たすことは複素関数が微分可能であることの必要条件である。微分可能な複素関数は \bar{z} に依存せず z だけの関数であることを示せ。

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

$f(z) = z^2 + 2z$ is a function of a complex number $z = x + iy$, where x and y are real numbers and i is the imaginary unit. Answer the following questions.

Q.1 Calculate the real part $u(x, y)$ and the imaginary part $v(x, y)$ of $f(z)$.

Q.2 Prove that $u(x, y)$ and $v(x, y)$ satisfy the following Cauchy-Riemann differential equations:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

Q.3 Prove that a function $g(\bar{z}) = \bar{z}^2 + 2\bar{z}$ of a complex number $\bar{z} = x - iy$ does not satisfy the Cauchy-Riemann differential equations.

Q.4 Satisfying the Cauchy-Riemann differential equations is a necessary condition for a complex function to be differentiable. Prove that a differentiable complex function is independent of \bar{z} and is a function of z only.

連結な無向グラフ G の部分グラフ G' で、頂点の集合がもとの G の頂点の集合と同じであり、かつ木であるものを G の全域木とよぶ。

設問 1 図 1 のグラフの全域木で頂点 1 と頂点 2 を結ぶ辺を含むものをすべて図示せよ。

頂点の集合を $V = \{1, 2, 3, 4\}$ とする完全グラフを K_4 とする。頂点の対 $(x, y) \in V \times V$ に対して、次の手順で K_4 の部分グラフ $H(x, y)$ を構成する。

Step 1 頂点 x と $z = \min(V - \{x, y\})$ を結ぶ。

Step 2 頂点 y と $w = \min(V - \{x, y, z\})$ を結ぶ。

Step 3 $V - \{z, w\}$ に含まれる 2 頂点を結ぶ。

なお、自然数の有限集合 S に対して $\min(S)$ は S の最小の要素を表す。

設問 2 $H(1, 3), H(2, 2)$ を図示せよ。

設問 3 任意の対 $(x, y) \in V \times V$ に対して、 $H(x, y)$ は K_4 の全域木であることを示せ。

設問 4 $(x_1, y_1) \neq (x_2, y_2)$ であれば $H(x_1, y_1) \neq H(x_2, y_2)$ であることを示せ。

アルファベットを $\Sigma = \{a, b\}$ とするとき、図 2 のような有限状態オートマトン M を考える。開始状態は p 、終了状態は q, s である。文字列 $w \in \Sigma^*$ が M によって受理されるとき $\delta(w) = 1$ 、受理されないとき $\delta(w) = 0$ と表すことにする。また、2 つの文字列 $u, v \in \Sigma^*$ を連接して得られる文字列を uv と表すことにする。

設問 5 $\delta(\epsilon), \delta(a), \delta(b), \delta(ab)$ の値をそれぞれ求めよ。ただし ϵ は空文字列を表す。

設問 6 $\delta(\epsilon w) = 0, \delta(aw) = 1, \delta(bw) = 1, \delta(abw) = 0$ の 4 つの条件をすべて満たす文字列 $w \in \Sigma^*$ を一つ与えよ。

設問 7 一般に有限状態オートマトン N が受理する文字列全体の集合を $L(N)$ と表す。図 2 の M に対して $L(M) = L(N)$ となる有限状態オートマトン N の状態数は 4 以上である。その理由を設問 5、設問 6 の結果を用いて説明せよ。

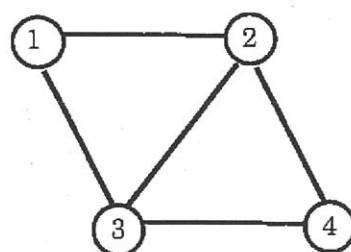


図 1

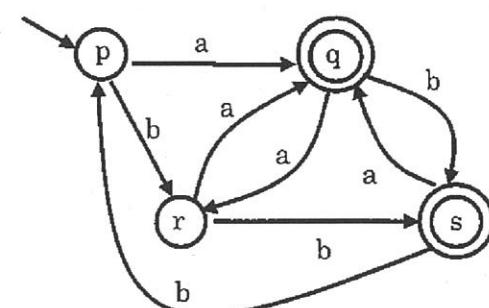


図 2

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

Let G be an undirected connected graph and G' be a subgraph of G . If G' is a tree and the set of vertices of G' is equal to the set of vertices of G , G' is said to be a spanning tree of G .

Q. 1 Draw all the spanning trees of the graph illustrated in Figure 1 in which vertex 1 and vertex 2 are connected by an edge.

Let K_4 be the complete graph of four vertices $V = \{1, 2, 3, 4\}$. For every pair of vertices $(x, y) \in V \times V$, the subgraph $H(x, y)$ of K_4 is defined by the following procedure:

- Step 1 Connect x and $z = \min(V - \{x, y\})$.
- Step 2 Connect y and $w = \min(V - \{x, y, z\})$.
- Step 3 Connect the two vertices in $V - \{z, w\}$.

Note that for a finite set S of distinct natural numbers, $\min(S)$ denotes the minimum element in S .

Q. 2 Draw $H(1, 3)$ and $H(2, 2)$.

Q. 3 Show that $H(x, y)$ is a spanning tree of K_4 for every pair $(x, y) \in V \times V$.

Q. 4 Show that if $(x_1, y_1) \neq (x_2, y_2)$ then $H(x_1, y_1) \neq H(x_2, y_2)$.

Let $\Sigma = \{a, b\}$ be an alphabet and consider the finite state automaton M illustrated in Figure 2. The initial state of M is p and its final states are q and s . For a string $w \in \Sigma^*$, we write $\delta(w) = 1$ if w is accepted by M and $\delta(w) = 0$ otherwise. For any two strings u and $v \in \Sigma^*$, uv denotes the string obtained by concatenating u and v .

Q. 5 Give the values of $\delta(\epsilon)$, $\delta(a)$, $\delta(b)$, and $\delta(ab)$, where ϵ is the empty string.

Q. 6 Find a string $w \in \Sigma^*$ which satisfies all of $\delta(\epsilon w) = 0$, $\delta(aw) = 1$, $\delta(bw) = 1$, and $\delta(abw) = 0$.

Q. 7 For any finite state automaton N , $L(N)$ denotes the set of all strings accepted by N . By using the answers to Q. 5 and Q. 6, explain the reason why any finite state automaton N satisfying $L(M) = L(N)$ has at least four states in it.

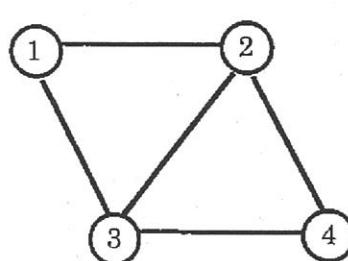


Figure 1

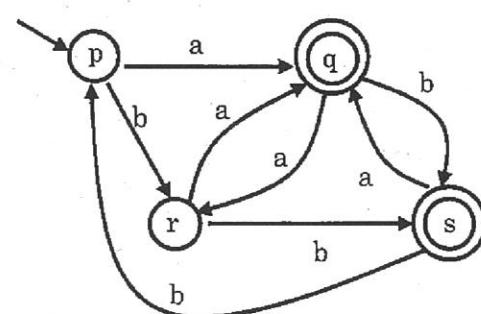


Figure 2

コインを投げたときに表が出る確率を $p \in [0, 1]$, 裏が出る確率を $1 - p$ とする。また、コインを 8 回投げたときに表が出る回数を k とする。

設問 1 k が 0 もしくは 8 となる確率 $P(k \in \{0, 8\})$ を、 $p = \frac{1}{2}$ および $p = \frac{1}{3}$ のそれぞれの場合について求めよ。

設問 2 p が未知のコインを考える。このコインが公正 ($p = \frac{1}{2}$) であるか否かを調べるために、回数 k を用いて、帰無仮説 $H_0: p = \frac{1}{2}$, 対立仮説 $H_1: p \neq \frac{1}{2}$ による仮説検定を行うことにする。ここで、棄却域 R_c を

$$R_c = \{0, \dots, c\} \cup \{8 - c, \dots, 8\}$$

(ただし $c \in \{0, 1, 2, 3, 4\}$) のように選ぶものとし、 $k \in R_c$ ならば H_0 を棄却する。このとき第一種の誤り確率 $P(k \in R_c | H_0 \text{ is true})$ が 0.15 未満になる c をすべて求めよ。

設問 3 設問 2 で求めた c による棄却域のうち、第二種の誤り確率 $P(k \notin R_c | H_0 \text{ is false})$ が最も小さくなるものはどれか説明せよ。

設問 4 いま、このコインが不公正に設計されており $p = \frac{1}{3}$ の可能性があるとする。この可能性を調べるために、回数 k を用いて、帰無仮説 $H_0: p = \frac{1}{3}$, 対立仮説 $H_1: p \neq \frac{1}{3}$ による仮説検定を行うことにする。ここで、棄却域 $R \subseteq \{0, 1, \dots, 8\}$ は任意に決められるものとしたとき、第一種の誤り確率が 0.15 未満になる棄却域のうち、第二種の誤り確率を最も小さくできる棄却域 R を求めよ。

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

Let $p \in [0, 1]$ be the probability of getting heads and $1 - p$ be the probability of getting tails when tossing a coin. Let k be the number of getting heads in eight tosses of a coin.

- Q.1 Calculate $P(k \in \{0, 8\})$ for the cases $p = \frac{1}{2}$ and $p = \frac{1}{3}$, respectively. Here $P(k \in \{0, 8\})$ is the probability that k becomes either 0 or 8.
- Q.2 Consider a coin with unknown p . To investigate whether the coin is fair ($p = \frac{1}{2}$) or not, let us consider conducting a hypothesis test with the null hypothesis $H_0: p = \frac{1}{2}$ and the alternative hypothesis $H_1: p \neq \frac{1}{2}$ by using k . Let R_c be the region of rejection chosen as
- $$R_c = \{0, \dots, c\} \cup \{8 - c, \dots, 8\},$$
- where $c \in \{0, 1, 2, 3, 4\}$; here H_0 is rejected when $k \in R_c$. Enumerate all c such that the probability of Type I error, $P(k \in R_c | H_0 \text{ is true})$, is below 0.15.
- Q.3 Explain which region minimizes the probability of Type II error, $P(k \notin R_c | H_0 \text{ is false})$, among the regions determined by c answered in Q.2.
- Q.4 Suppose there is a possibility that the coin has a designed bias $p = \frac{1}{3}$. To investigate this possibility, let us consider conducting a hypothesis test with the null hypothesis $H_0: p = \frac{1}{3}$ and the alternative hypothesis $H_1: p \neq \frac{1}{3}$ by using k . Suppose now that the region of rejection $R \subseteq \{0, 1, \dots, 8\}$ can be decided arbitrarily. Find the region R for the test that minimizes the probability of Type II error while making the probability of Type I error less than 0.15.