

平成 23 年度 10 月期入学 / 平成 24 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 23 年 8 月 8 日 13:00~16:00

【注意】

1. 問題冊子はこの表紙を含めて 39 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 25 題であり、一部の問題は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-8)	1-8 ページ
心理学(問題番号:P-1~P-4)	9-12 ページ
計算機科学, 電気電子工学(問題番号:T-1~T-13)	13-38 ページ

5. 特に指定のない限り、日本語または英語で解答しなさい。
6. 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。

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

October 2011 Admissions / April 2012 Admissions
Entrance Examination for Master's Program
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
(Specialized Subjects)

August 8, 2011
13:00 - 16:00

NOTES

1. This is the Problem Booklet in 39 pages including this front cover.
2. Do not open the pages until you are instructed to start.
3. After start, check all pages in order and notify proctors (professors) immediately if missing pages or unclear printings are found.
4. There are 25 problems, and some problems are written in Japanese and English. The problems are classified as listed below. **Choose and solve 4 problems.**

Bioinformatics (Problem Number B-1 to B-8)	Page 1 to 8	
Psychology (Problem Number P-1 to P-4)	Page 9 to 12	
Computer science, Electrical and electronic engineering (Problem Number T-1 to T-13)		Page 13 to 38

5. Solve the problems in Japanese or English, unless otherwise specified.
6. Read carefully the notes on the front cover of the Answer Sheets as well.

これはある論文の要約と図の一部である。以下の設問に答えよ。



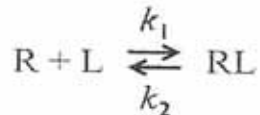
No Disclosure

(Ofstad et al., Nature, 474, 204-207, 2011 から改変).

- 設問 1 下線部 A, B を日本語に翻訳せよ。
- 設問 2 この実験に書かれている実験結果を図に示す。この図から得られることを理由とともに述べよ。
- 設問 3 ネズミを使った類似の実験がある。どのような実験か述べよ。ネズミの実験と本実験を比較し、本実験の欠点、利点をのべよ。
- 設問 4 この論文にタイトルをつけよ(日本語でも英語でもよい)。

以下の文章を読み設問に答えよ。

細胞膜(plasma membrane)表面上の受容体(receptor)R とリガンド(ligand)L の結合について、リガンドと受容体の結合の速度定数(rate constant)を k_1 、解離の速度定数を k_2 とすると以下のように表すことができる。



設問1. 受容体とリガンドの結合を測定する方法を述べよ。

設問2. 受容体の濃度(concentration)が $[R]$ のタンパク質溶液に、時刻 $t = 0$ のとき濃度 $[L] = L_0$

のリガンドを加えた。リガンドが結合した受容体の濃度 $[RL]$ の時間変化 $\frac{d[RL]}{dt}$ はどのように表されるか述べよ。

設問3. 設問1の状態から十分時間がたち平衡(equilibrium)な状態に達したとき、見かけ上リガンドが結合した受容体は何%になるか答えよ。

設問4. 平衡状態にあるひとつの受容体に着目する。時刻 0 から t までリガンドと結合している確率を $P(t)$ とする。時刻 t からさらに Δt 後もリガンドと結合している確率 $P(t+\Delta t)$ はどのように表すことができるか説明せよ。

設問5. 平衡状態にあるひとつの受容体に着目する。リガンドと結合している平均時間 T を求めよ。

以下の文章を読み、設問に答えよ。

神経筋接合部(neuromuscular junction)では、運動神経(motor neuron)からアセチルコリン(acetylcholine)が放出され筋肉(muscle)の収縮を引き起こしている。アセチルコリンは筋肉で終板電位(end-plate potential)を発生させる。筋肉の終板電位を測定したところ約 1mV の脱分極(depolarization)が自発的(spontaneous)に不規則に起こっていた。自発性終板電位を多数観測し、振幅(amplitude)のヒストグラムを作成した(図 1 A)。さらに解析をすすめるために細胞外液(extracellular solution)のカルシウム(calcium)濃度を低下させ、神経刺激で誘発した終板電位を多数測定し、振幅のヒストグラムを作成した(図 1 B)。

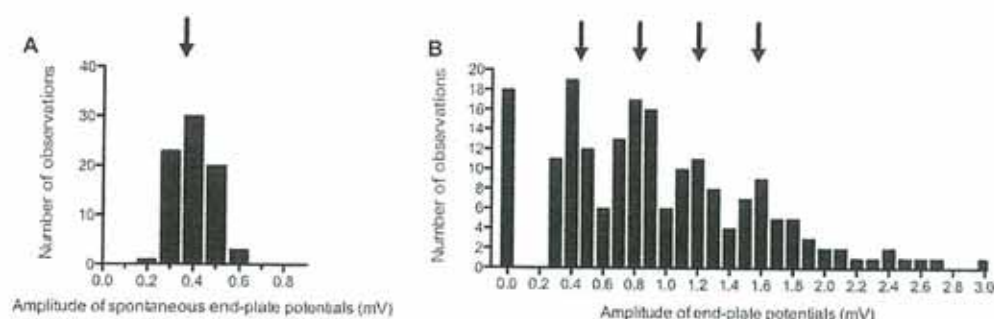


図 1 自発性の終板電位の分布 (A) と神経刺激で誘発された終板電位の分布 (B)。矢印(arrow)は観測数のピークを示す。

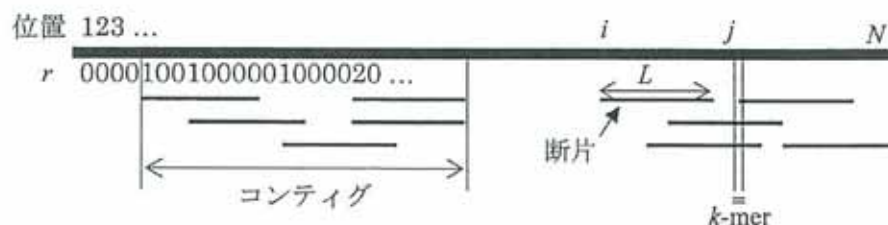
- 設問1. 神経刺激で誘発された終板電位は、自発性終板電位の平均振幅の 1 倍、2 倍、3 倍、4 倍のところに観測数のピークがみられる(図 1B 矢印)。このことから神経刺激により誘発される終板電位とアセチルコリンの量との関係を考察せよ。
- 設問2. 細胞外液にアセチルコリン受容体のアンタゴニスト(antagonist)をさまざまな濃度で投与し、自発性終板電位を測定した。アンタゴニストの濃度を上げていくと、自発性終板電位の振幅は連続的に小さくなるが頻度は変わらなかった。このことから、自発性終板電位はアセチルコリンが一定量放出されたことによると考えられるが、その根拠を述べよ。
- 設問3. 細胞外液に少量のアセチルコリンを加えた場合、自発性電位の頻度はどのように変化すると考えられるか、理由とともに述べよ。
- 設問4. ヒトやマウスでは、アセチルコリンは神経筋接合部以外に脳でも様々な役割をになっていることが知られている。その役割をひとつ述べ、説明せよ。

以下の語句の中から5つを選び、簡単に説明せよ。

1. 沈降平衡法 (sedimentation equilibrium)
2. 高速液体クロマトグラフィー (high-performance liquid chromatography: HPLC)
3. 等電点電気泳動 (isoelectric focusing)
4. 遺伝子すりこみ (genomic imprinting)
5. アロステリック制御 (allosteric regulation)
6. 核磁気共鳴 (nuclear magnetic resonance)
7. 共焦点レーザー走査型顕微鏡 (confocal laser scanning microscopy)
8. アルファヘリックス (alpha helix)
9. ファンデルワールス力 (van der Waals force)
10. ゴルジ体 (Golgi apparatus)

全長 N のゲノム DNA 塩基配列 (genomic DNA sequence) $G=g_1g_2\ldots g_N$, $g_i \in \{A, C, G, T\}$ の配列決定実験 (sequencing) を考える。1 回の試行 (trial) で、区間 $[1, N-L+1]$ から無作為 (at random) に選んだ位置 i を始点 (start point) とする長さ L の断片 (fragment) $g_i g_{i+1} \ldots g_{i+L-1}$ の塩基配列を読む (決定する) ことができ、この試行を R 回繰り返すものとする。ここで $L \ll N$ であり、端の効果 (end effects) は無視できるものとする。このとき以下の設問に答えよ。

- 設問 1 G 上の特定の位置 i の塩基 g_i が r 個の断片の始点である確率は $P(r; a) = \frac{a^r e^{-a}}{r!}$ の形で表すことができる。 a を既知の変数を用いて表せ。この確率分布の名称を答えよ。また、 r の期待値 (expected value) を示せ。
- 設問 2 N カ所の位置のうち、少なくとも 1 つの断片に含まれるものの割合 (既読率、read fraction) の期待値を求めよ。
- 設問 3 内部に未読の塩基を含まない連続した既読領域 (read region) をコンティグ (contig) とよぶ (下図参照)。コンティグ数の期待値を求めよ。隣り合う断片間の距離が $L+1$ 以上であるとコンティグが途絶えることに注意せよ。
- 設問 4 長さ $k (< L)$ の部分塩基配列 (substring) を k -mer とよぶ。 G 上の k -mer の数を示せ。また、 R 個の断片に存在する k -mer の総数を示せ。相補鎖 (complementary strand) を考慮する必要はない。
- 設問 5 任意に選んだ G 上の位置 j を始点とする k -mer が、 x_j 個 (下図では 2 個) の断片に含まれていたとする。そのような事象の起きる確率を求めよ。これを利用して未知のゲノム長 N を推定する方法について述べよ。ただし、 G 上の塩基配列はランダム配列であると仮定してよい。また、 $N < 4^k$ であり、同じ k -mer が G 上の複数の位置に現れる可能性は無視できるものとする。



太い実線はゲノムを、細い実線は塩基配列の読まれた各断片を、二重線は k -mer を表す。

置換 (substitution) の蓄積によって共通の祖先配列から分岐 (divergence) した 2 本の DNA 配列について、その一致度 (identity) Q を考える。置換とは、配列中のある座位 (site) の塩基 (base) が別の 3 種類の塩基のどれかに置き換わることを指す。一致度は、2 本の DNA 配列の塩基が一致する割合で、配列 ATCTTGC と AACCTGC の場合は 5/7 になる (下図)。

ATCTTGC
: : : :
AACCTGC

(1) 置換の起きる確率は配列中の全ての座位で等しく、(2) どの塩基からどの塩基への置換も等しい確率で起きると仮定する。また、DNA 配列は十分に長いものとする。以下の設問に答えよ。

設問 1 座位あたり単位時間あたりの置換速度 (substitution rate) を r とすると、一致度 Q の時間変化 $\frac{dQ}{dt}$ が次式で表わされることを示せ。

$$\frac{dQ}{dt} = \frac{2r}{3} - \frac{8r}{3}Q$$

ここで、 r は非常に小さいので、 r^2 の項は無視できるとする。

設問 2 時刻 $t = 0$ における一致度を $Q_0 = 1$ として、上記の微分方程式 (differential equation) を解け。

設問 3 設問 2 の解を変形して、2 本の DNA 配列の一致度 Q からそれらの進化距離 (evolutionary distance) $2rt$ を計算する式を示せ。

設問 4 設問 2 の解より、時間が十分に経過すると DNA 配列の塩基組成 (base composition) はどのようなになるか、述べよ。

設問 5 上記の (1) と (2) の仮定は、実際には成り立たないことが多い。各々について、成り立たない場合の具体例を述べよ。

生命情報学においては生物情報ネットワーク (biological information network) をグラフ (graph) としてモデル化して, その構造上の特徴を調べる研究が行われている. 以下ではグラフとして, 多重辺 (multi-edge) なし, かつ, 自己ループ (self-loop) なしの連結 (connected) な無向グラフ (undirected graph) $G(V, E)$ のみを対象として議論する. ただし, V は頂点集合 (set of vertices), E は辺集合 (set of edges) である. ここで, 頂点 v の次数 (degree) $d(v)$, および, クラスター係数 (clustering coefficient) $C(v)$ (ただし, $d(v) > 1$ の頂点のみを対象とする) を次のように定義する.

$$d(v) = |\{w \mid \{v, w\} \in E\}|$$

$$C(v) = \frac{2 \cdot |\{\{u, w\} \mid u \neq w, \{u, v\} \in E, \{w, v\} \in E\}|}{d(v)(d(v) - 1)}$$

つまり, $d(v)$ は v に隣接する頂点の個数であり, $C(v)$ は v に隣接する頂点間における辺の頻度を表す. たとえば, 下の図のグラフにおいて

$$\begin{aligned} d(v_1) &= 2 & d(v_2) &= 4 & d(v_3) &= 2 & d(v_4) &= 3 \\ C(v_1) &= 1 & C(v_2) &= 1/2 & C(v_3) &= 1 & C(v_4) &= 2/3 \end{aligned}$$

となる. 以下の設問に答えよ.

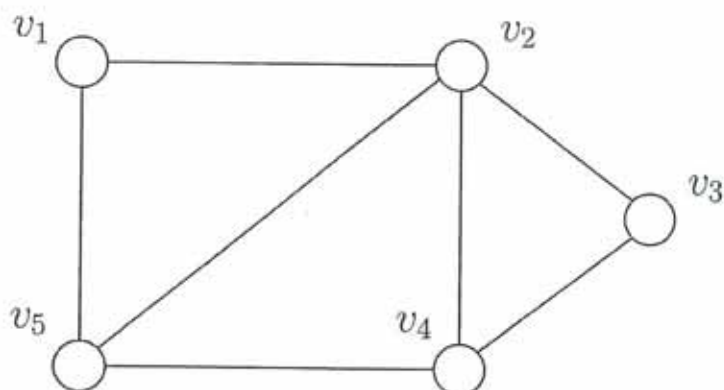
設問 1 生物情報ネットワークの例 (種類) を一つあげ, 頂点は何に, 辺が何に対応するか述べよ.

設問 2 下の図のグラフにおける $d(v_5)$ と $C(v_5)$ を求めよ.

設問 3 以下のそれぞれについて条件を満たすグラフが存在する場合にはそのグラフを示し, 存在しない場合にはその理由を示せ.

- (A) $|V| = 4$ で, すべての頂点 v に対し $d(v) = 3$, かつ, $C(v) = 1$ が成立.
- (B) $|V| = 8$ で, すべての頂点 v に対し $d(v) = 3$, かつ, $C(v) = 1$ が成立.
- (C) $|V| = 6$ で, すべての頂点 v に対し $d(v) = 3$, かつ, $C(v) = 0$ が成立.

設問 4 $|V| \geq 10000$ を満たし, すべての頂点 v に対し $d(v) = 3$, かつ, $C(v) = 0$ が成立するグラフが存在するかどうかを考え, 存在する場合にはその一つの具体的構成法 (construction method) を示し, 存在しない場合にはその理由を明確に示せ.



生命情報学に関する以下の語句の中から 8 個を選んで説明せよ。必要に応じて図を用いて説明してもよい。

1. メタゲノム (metagenome)
2. ヘリックスバンドル (helix bundle)
3. GenBank
4. MEDLINE
5. マルチプルアラインメント (multiple alignment)
6. 近隣結合法 (neighbor-joining method)
7. コンタクトマップ (contact map)
8. RNA 二次構造 (RNA secondary structure)
9. ネットワークモチーフ (network motif)
10. 代謝流束解析 (metabolic flux analysis)

以下の文を読んで設問に答えよ。

A, B, C, D の4つの文字が、図1に示される状態遷移図(state transition diagram)に従って、画面に一文字ずつ継時的に提示される(sequentially presented)。このとき、実験協力者(participant)には、各文字に対応する反応キーをできるだけはやく、かつ正確に押すことが求められた。このような試行を多数回行くと、徐々に協力者の反応時間(reaction time)が短縮することが分かる。しかし、協力者自身は、この4文字の系列(sequence)がどのような規則に従って生成されているかは答えることができない。

設問1 このような実験は、一般に人工文法(artificial grammar)の学習実験と呼ばれる。なぜか。

設問2 協力者は、文字系列の生成規則(rule of generation)を言えないにも関わらず反応時間が短縮した。この理由について考察せよ。

設問3 人工文法の学習実験では、潜在的に(implicitly)文字系列の学習が進められているかどうかを客観的に調べるある方法が使われる。どのような方法かを述べよ。

設問4 内側側頭葉性健忘症(mesial temporal amnesia)の患者(patient)でもこのような学習が可能であるということが知られている。これは記憶研究にとって大変重要な発見(finding)なのであるが、その理由について述べよ。

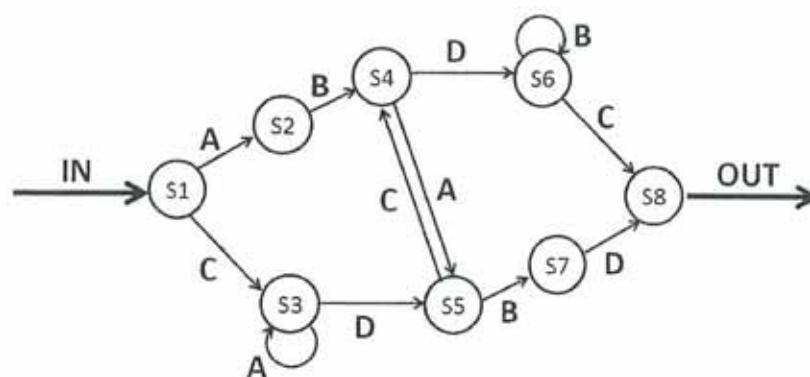


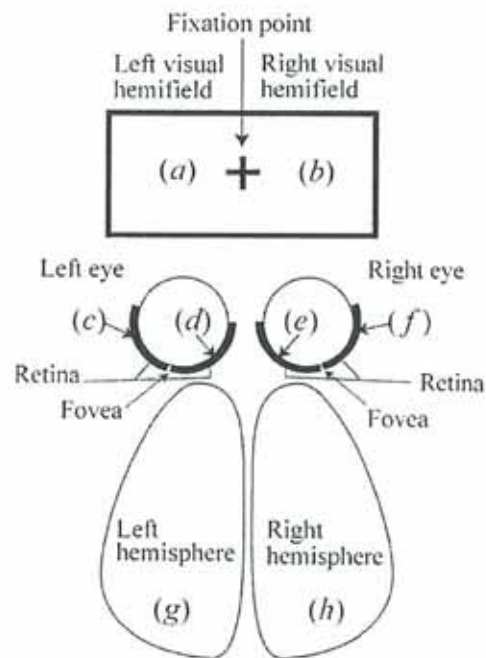
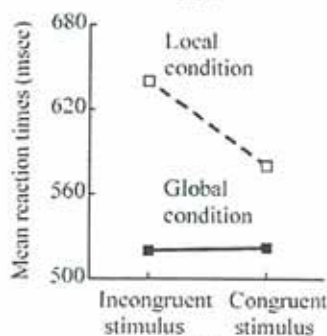
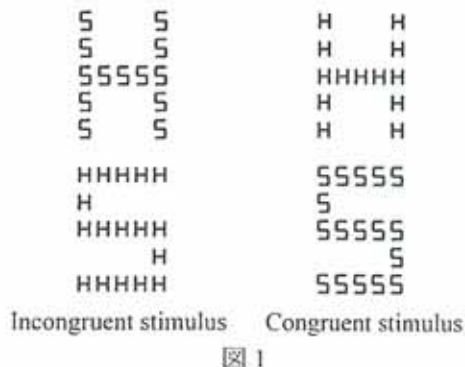
図1

図1左に示す2種類の「不一致刺激」(incongruent stimulus)と、図1右に示す2種類の「一致刺激」(congruent stimulus)の内、いずれかを1つずつ実験協力者(participant)に提示する実験を行った。課題(task)は2種類あり、「局所条件」(local condition)では、実験協力者は小さい文字の名前を答え、「大域条件」(global condition)では小さい文字が何であるかに関わらず、小さい文字で構成された大きい文字の名前を答えた。以下の設問に答えよ。

設問1 図2は不一致刺激、一致刺激に対する平均反応時間(mean reaction time)である。分散分析(analysis of variance)の結果、課題(大域条件・局所条件)の主効果(main effect)、刺激の種類(一致・不一致)と課題の交互作用(interaction)が有意であった。この結果から、大域的な情報と局所的な情報がどのように処理されているか考察せよ。

設問2 左または右の視野(visual hemifield)だけに図1の刺激を提示する実験を考える。このとき、刺激は網膜(retina)を通して大脳の左半球(left hemisphere)や右半球(right hemisphere)へと伝達(transmit)される。このときの情報の伝達経路として考えられるものを、図3中の(a)~(h)を用いて“(a)→(x)→(y)”のようにして全て示せ。ただし、実験協力者は固視点(fixation point)を固視しているものとし、脳梁(corpus callosum)を介した伝達は考えない。

設問3 大脳の左半球は分析的(analytic)な情報処理を行い、右半球は全体的(holistic)な情報処理を行うという仮説がある。この仮説が正しいとすれば、設問2のような実験を行ったときの反応時間は、課題(大域条件・局所条件)と視野(左視野・右視野)に依存してどのように変化すると予測されるか。脳梁を介した情報伝達も考慮して考察せよ。



以下の文章を読んで設問に答えよ。

実験参加者(participant)には cue 刺激(cue stimulus) (●)が固視点(fixation point)の左右視空間(left or right hemifield)のいずれに提示されたかをボタン押し(key press)で回答するように教示(instruction)した。なお, cue 刺激に先立ち, precue 刺激(□)を cue 刺激が提示される視空間と同側(ipsilateral hemifield), または反対側(contralateral hemifield)に提示した。このとき, cue 刺激提示から実験参加者によるボタン押しまでの反応時間(reaction time)を測定した。

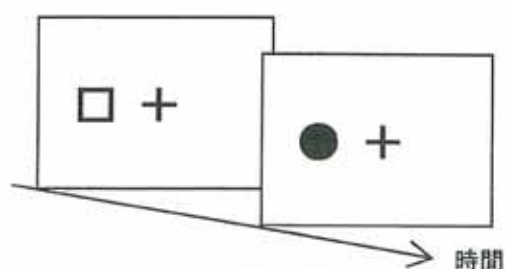


図 1 実験課題

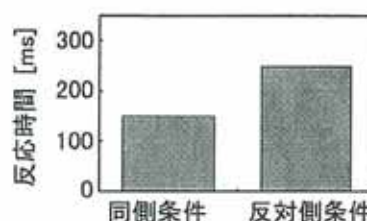


図 2 反応時間の測定結果

設問 1 ある実験参加者から反応時間を測定したところ, 表 1 に示すデータを得た。この反応時間の代表値(central tendency)を求めよ。また, その導出根拠(reason)を簡潔に述べよ。なお, $\log_{10}2=0.301$, $\log_{10}3=0.477$, $\log_{10}5=0.699$ を用いてもよい。

表 1 反応時間の測定結果

反応時間[ms]	100	200	200	200	200	300	300	400	500	800
----------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

設問 2 precue 刺激と同じ視空間位置に cue 刺激が提示される同側条件と, 反対側に提示される反対側条件を実施したとき, 図 2 に示すような反応時間の測定結果を得た。図 2 に示すような反応時間の差が発生する理由を述べよ。

設問 3 precue 刺激を提示せずに cue 刺激のみを提示する実験を行った場合, 図 2 に示す同側条件・反対側条件の反応時間の測定結果と比較して, 反応時間はどのように変化するか。その理由とともに答えよ。

設問 4 複数の実験参加者に対して同様の実験を実施したとき, 同側条件・反対側条件での反応時間の差の検定には, 固定効果モデル(fixed effect model)と変量効果モデル(random effect model)を用いる方法が考えられる。固定効果モデルと変量効果モデルについて説明せよ。

以下の英文を読んで設問に答えよ。

No Disclosure

(Parsons (1987) *Cognitive Psychology*, 19 178-241. より改変)

設問 1 この実験における独立変数(independent variable)と従属変数(dependent variable)を述べよ。

設問 2 各実験参加者(participant)が従属変数の測定のために行った総試行数(total number of trials)を述べよ。

設問 3 実験参加者はどのような方略(strategy)で判断をしたと報告しているかを述べよ。また、その内容に基づき back 刺激において従属変数が独立変数の単調増加関数(monotonically increasing function)となる理由を考察せよ。

設問 4 palm 刺激において lateral orientation と medial orientation で結果に大きな違いが生じた理由を考察せよ。

The program called "sorting.sh" as shown in Fig. 1 is a simple sorting Bash program that accepts a sequence of positive integers. `sleep` is a Unix command that suspends thread execution for an interval measured in seconds. The unit time of the internal clock of the system is sufficiently small compared to a second.

Answer the following questions.

Question 1 Suppose that there is no overhead in executing `sorting.sh`; that is, any Bash command can be executed without delay. When the following command is executed,

`sorting.sh 30 12 1 5 8 12 3 23`

in what order is each data given in the output? Fill in the following table in the order of execution time.

Time after the command (sec)								
output data								

File name: `sorting.sh`

```
#!/bin/bash
function f() {
    sleep "$1"
    echo "$1"
}
while [ -n "$1" ]
do
    f "$1" &
    shift
done
wait
```

Fig. 1. Sorting in Bash
[Anonymous 2011]

Question 2 Explain whether this sorting program is stable or not, and why.

Question 3 Suppose that there is no overhead in executing `sorting.sh`. Explain the latency after running the program.

Question 4 Suppose that the overhead in executing `sorting.sh` is not zero. Explain the maximum overhead time makes the program work correctly.

Question 5 Design the data structure and its operations to maintain threads that makes `sleep` work. You can use any language.

Question 6 Describe the time and space complexities of the operations answered in Question 5.

Question 7 Describe a modification that will speed up the execution of the sorting program specified in Fig. 1.

図1に示した“sorting.sh”は、Bashによる正整数の系列の簡単な整列プログラムである。sleepは指定された時間(秒)だけ thread の実行を遅延する Unix コマンドである。システムの内部時計の単位は秒よりも十分に小さいものとする。

この時、以下の質問に答えよ。

設問1 sorting.shの実行に全くオーバーヘッドがないとする。すなわち、いかなる Bash コマンドも遅延なしで実行される。次のコマンドが実行されると、

sorting.sh 30 12 1 5 8 12 3 23

どのようなデータがどのような順序で出力されるか、次の表を埋めよ。

コマンド入力後の経過時間 (sec)									
出力データ									

File name: sorting.sh

```
#!/bin/bash
function f() {
    sleep "$1"
    echo "$1"
}
while [ -n "$1" ]
do
    f "$1" &
    shift
done
wait
```

図1. Bash の整列プログラム
[Anonymous 2011]

設問2 本整列プログラムは stable かどうかを理由と共に述べよ。

設問3 sorting.shの実行に全くオーバーヘッドがないとき、本整列プログラムが完了するまでの遅延時間を述べよ。

設問4 オーバヘッドが0でないとすると、本プログラムが正しく働くための最大のオーバーヘッド時間について述べよ。

設問5 sleepを実現するために thread を管理するのに使用するデータ構造とその演算を設計せよ。どのような言語を用いてもよい。

設問6 設問5で設計した演算の時間・空間計算量について述べよ。

設問7 図1に示した整列プログラムの実行を高速化するための修正について述べよ。

自然数全体からなる集合 N と正の整数 k に対して,

$$N^k = \{(n_1, \dots, n_k) \mid n_i \in N, i = 1, 2, \dots, k\}$$

とし, N^k の要素をタプル, N^k の有限部分集合を k 次元データベースとよぶことにする. 以下では, 正の整数 k, h が $k > h > 0$ を満たすとし, R と S をそれぞれ k 次元データベース, h 次元データベースとする.

演算 $\times, \div, \%$ をそれぞれ次のように定義する.

$$R \times S = \{(n_1, \dots, n_k, m_1, \dots, m_h) \mid (n_1, \dots, n_k) \in R \text{ かつ } (m_1, \dots, m_h) \in S\},$$

$$R \div S = \left\{ (n_1, \dots, n_{k-h}) \mid \begin{array}{l} \text{任意の } (m_1, \dots, m_h) \in S \text{ に対して} \\ (m_1, \dots, m_h, n_1, \dots, n_{k-h}) \in R \end{array} \right\},$$

$$R \% S = R - (S \times (R \div S)).$$

ただし集合 A と B の集合差を $A - B$ で表している. 以下の設問に答えよ.

設問1 次の4次元データベース R_1 と2次元データベース S_1 に対して, $R_1 \% S_1$ を求めよ.

$$R_1 = \left\{ \begin{array}{ll} (1, 4, 7, 9), & (2, 5, 7, 9), \\ (3, 6, 8, 9), & (1, 5, 8, 9), \\ (3, 6, 7, 9), & (1, 4, 8, 9), \\ (2, 5, 8, 9), & (1, 5, 7, 9) \end{array} \right\}, \quad S_1 = \left\{ \begin{array}{l} (1, 4), \\ (1, 5), \\ (3, 6) \end{array} \right\}.$$

設問2 等式 $(S \times R) \div S = R$ が成立することを証明せよ.

設問3 タプル $t = (m_1, \dots, m_h)$ に対して, $R \div \{t\}$ を計算するアルゴリズム $\text{unitdiv}(R, t)$ を, 日本語または英語で与えよ.

設問4 アルゴリズム $\text{unitdiv}(R, t)$ を効率よく実行するためには, R の要素をどのようなデータ構造で保持すればよいのか?

設問5 データベースに対する集合演算と $\text{unitdiv}(R, t)$ を用いて $R \div S$ を計算するアルゴリズム $\text{div}(R, S)$ を, 日本語または英語で与えよ.

次に, データベース R が与えられたとき,

$$R \% S = \emptyset \tag{1}$$

を満たす S を求めることを考える. なお, \emptyset は空集合を表す. 以下の設問に解答せよ.

設問6 設問1のデータベース R_1 に対して $R = R_1$ としたときに, 等式 (1) を満たす S を $h = 1, 2, 3$ の各次元について, それぞれ1つずつ挙げよ. なお次元 h によっては, そのような S が存在しないことがある.

設問7 一般に, 任意の R に対して, 等式 (1) を満たす S は各次元 h ($k > h > 0$) について高々1つしか存在しないことを証明せよ.

Let \mathbf{N} be the set of natural numbers. For a positive integer k , we define

$$\mathbf{N}^k = \{(n_1, \dots, n_k) \mid n_i \in \mathbf{N}, i = 1, 2, \dots, k\},$$

then we call every element of \mathbf{N}^k a *tuple* and every finite subset of \mathbf{N}^k a *database* of k -dimension. From now we assume that two integers k and h satisfy $k > h > 0$, and R and S are respectively databases of k -dimension and h -dimension.

We define operators \times , \div , and $\%$ as follows:

$$\begin{aligned} R \times S &= \left\{ (n_1, \dots, n_k, m_1, \dots, m_h) \mid (n_1, \dots, n_k) \in R \text{ and } (m_1, \dots, m_h) \in S \right\}, \\ R \div S &= \left\{ (n_1, \dots, n_{k-h}) \mid \begin{array}{l} \text{For all } (m_1, \dots, m_h) \in S \\ (m_1, \dots, m_h, n_1, \dots, n_{k-h}) \in R \end{array} \right\}, \\ R \% S &= R - (S \times (R \div S)), \end{aligned}$$

where $A - B$ denotes the difference of two sets A and B .

Answer the following questions.

Question 1 Give out $R_1 \% S_1$ where R_1 and S_1 are respectively databases of 4-dimension and 2-dimension provided as:

$$R_1 = \left\{ \begin{array}{ll} (1, 4, 7, 9), & (2, 5, 7, 9), \\ (3, 6, 8, 9), & (1, 5, 8, 9), \\ (3, 6, 7, 9), & (1, 4, 8, 9), \\ (2, 5, 8, 9), & (1, 5, 7, 9) \end{array} \right\}, \quad S_1 = \left\{ \begin{array}{l} (1, 4), \\ (1, 5), \\ (3, 6) \end{array} \right\}.$$

Question 2 Prove that the equation $(S \times R) \div S = R$ is true.

Question 3 For a tuple $t = (m_1, \dots, m_h)$, give out an algorithm *unitdiv*(R, t) in Japanese or in English which takes R and t as its inputs, and outputs $R \div \{t\}$.

Question 4 In order to make the execution of *unitdiv*(R, t) efficient, what data structure should R be stored in?

Question 5 Give out an algorithm *div*(R, S) which computes $R \div S$ by using some set operations on databases and the algorithm *unitdiv*(R, t).

Let us consider to find S for given R such that

$$R \% S = \emptyset, \tag{1}$$

where \emptyset denotes an empty set. Answer the following questions.

Question 6 Assume that $R = R_1$ where R_1 is the database provided in Question 1. For each of the dimensions $h = 1, 2, 3$, find an S satisfying the equation (1). Note that S may not always exist in these cases.

Question 7 For an arbitrary R , prove that there exists at most one S satisfying the equation (1) for each dimension h ($k > h > 0$).

論理式の集合 Fm を以下の BNF (Backus-Naur form) で定義する.

$$Fm ::= PV \mid \perp \mid \neg Fm \mid Fm \wedge Fm$$

ここで, PV は命題変数の集合とする. 以下では p, q などで PV の要素を, A, B などで Fm の要素を表すとする. $A \rightarrow B$ は $\neg(A \wedge \neg B)$ の略記とする.

論理体系 ND を以下の推論規則で定義する. ここで, Γ は論理式の有限集合を表すとする.

$$\begin{array}{c} \frac{}{\Gamma \cup \{A\} \vdash A} \quad \frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad \frac{\Gamma \vdash \neg(\neg A)}{\Gamma \vdash A} \quad \frac{\Gamma \cup \{A\} \vdash \perp}{\Gamma \vdash \neg A} \quad \frac{\Gamma \vdash A \quad \Gamma \vdash \neg A}{\Gamma \vdash \perp} \\ \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} \end{array}$$

設問 1 以下の判断 1~3 の ND における導出を与えよ.

1. $\emptyset \vdash A \rightarrow A$
2. $\{A \rightarrow B, A\} \vdash B$
3. $\{(\neg B) \rightarrow \neg A\} \vdash A \rightarrow B$

設問 2 \mathbb{B} を集合 $\{0, 1\}$ とする. 写像 $\rho: PV \rightarrow \mathbb{B}$ に対して, $[A]_\rho \in \mathbb{B}$ を以下のように定義する.

$$[p]_\rho = \rho(p) \quad [\perp]_\rho = 0 \quad [\neg A]_\rho = 1 - [A]_\rho \quad [A \wedge B]_\rho = [A]_\rho \times [B]_\rho$$

ここで, $-$ と \times は通常自然数の減算と乗算である.

$\rho \models A$ によって, $[A]_\rho = 1$ であることを表し, 論理式の有限集合 Γ について, $\rho \models \Gamma$ によって, 全ての $B \in \Gamma$ が $\rho \models B$ を満たすことを表す. さらに, $\models A$ によって, 任意の $\rho: PV \rightarrow \mathbb{B}$ について $\rho \models A$ であることを表す.

1. $[\neg p]_\rho$ の値は表を用いて

$\rho(p)$	$[\neg p]_\rho$
1	0
0	1

と表せる. これにならって, $\rho(p)$ と $\rho(q)$ の値の全ての組合せに対する $[p \wedge q]_\rho$, $[p \rightarrow q]_\rho$ の値を表せ.

2. 任意の A について $\models A \rightarrow A$ が成り立つことを示せ.
3. $[(p \rightarrow q) \rightarrow (q \rightarrow p)]_\rho = 0$ となるような ρ が存在することを示せ.
4. $\rho \models \{A \rightarrow B, A\}$ なる任意の ρ について $\rho \models B$ が成り立つことを示せ.
5. $\Gamma \vdash A$ が ND で導出可能であるとき, 任意の $\rho: PV \rightarrow \mathbb{B}$ について, $\rho \models \Gamma$ ならば $\rho \models A$ であることを示せ.

The set of formulas Fm is defined by the following BNF (Backus-Naur form):

$$Fm ::= PV \mid \perp \mid \neg Fm \mid Fm \wedge Fm,$$

where PV is the set of propositional variables, denoted by p, q, \dots . Elements of Fm are denoted by A, B, \dots . We use an abbreviation $A \rightarrow B$ to denote $\neg(A \wedge \neg B)$.

The logical system ND is defined by the following inference rules, where Γ denotes a finite set of formulas.

$$\begin{array}{c}
 \frac{}{\Gamma \cup \{A\} \vdash A} \quad \frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad \frac{\Gamma \vdash \neg(\neg A)}{\Gamma \vdash A} \quad \frac{\Gamma \cup \{A\} \vdash \perp}{\Gamma \vdash \neg A} \quad \frac{\Gamma \vdash A \quad \Gamma \vdash \neg A}{\Gamma \vdash \perp} \\
 \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}
 \end{array}$$

Question 1 Give derivations of the following judgments in ND.

1. $\emptyset \vdash A \rightarrow A$
2. $\{A \rightarrow B, A\} \vdash B$
3. $\{(\neg B) \rightarrow \neg A\} \vdash A \rightarrow B$

Question 2 Let \mathbb{B} be the set $\{0, 1\}$. For a map $\rho : PV \rightarrow \mathbb{B}$, we define $\llbracket A \rrbracket_\rho \in \mathbb{B}$ as follows:

$$\llbracket p \rrbracket_\rho = \rho(p) \quad \llbracket \perp \rrbracket_\rho = 0 \quad \llbracket \neg A \rrbracket_\rho = 1 - \llbracket A \rrbracket_\rho \quad \llbracket A \wedge B \rrbracket_\rho = \llbracket A \rrbracket_\rho \times \llbracket B \rrbracket_\rho,$$

where $-$ and \times are the ordinary subtraction and multiplication on natural numbers, respectively.

$\rho \models A$ denotes that $\llbracket A \rrbracket_\rho = 1$ holds. For a finite set Γ of formulas, $\rho \models \Gamma$ denotes that $\rho \models B$ holds for any $B \in \Gamma$. $\models A$ denotes that $\rho \models A$ for any map $\rho : PV \rightarrow \mathbb{B}$.

1. The value of $\llbracket \neg p \rrbracket_\rho$ can be represented by the following table.

$\rho(p)$	$\llbracket \neg p \rrbracket_\rho$
1	0
0	1

Show the tables of the values of $\llbracket p \wedge q \rrbracket_\rho$ and $\llbracket p \rightarrow q \rrbracket_\rho$ for all combinations of the values of $\rho(p)$ and $\rho(q)$.

2. Show that $\models A \rightarrow A$ holds for any A .
3. Show that there exists a map ρ such that $\llbracket (p \rightarrow q) \rightarrow (q \rightarrow p) \rrbracket_\rho = 0$.
4. Show that $\rho \models B$ holds for any ρ such that $\rho \models \{A \rightarrow B, A\}$.
5. Show that if $\Gamma \vdash A$ is derivable in ND, then, for any $\rho : PV \rightarrow \mathbb{B}$, $\rho \models \Gamma$ implies $\rho \models A$.

記憶のない定常 2 元対称通信路 $T = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$ を通して N 種類のメッセージ $A = \{a_1, a_2, \dots, a_N\}$ を伝送するため, ランダム符号化を適用し, 送信側で各 $a_i \in A$ ($i=1, \dots, N$) に対してランダムに生成された長さ n の符号語 $x_1 x_2 \dots x_n$ ($x_j \in \{0, 1\}$) を割り当てるものとする. 受信側では各符号語を伝送するときに通信路で発生する誤りの個数が np 以下であると仮定して復号する. ただし, $p \leq \frac{1}{2}$ であり, n は十分大きいものとする.

設問 1 T の通信路容量 C を示せ.

設問 2 この通信路を使って各符号語を伝送するときに生じる誤りの個数を表す確率変数を Z とすると, $Z = k$ ($k = 0, 1, \dots, n$) となる確率は $P\{Z = k\} = {}_n C_k p^k (1-p)^{n-k}$ である. ランダム符号化における Z の期待値 $E(Z)$ は np であることを示せ.

設問 3 与えられた符号語に対して, ハミング距離が np 以下となる長さ n の 2 元系列

の個数を $M = \sum_{k=0}^{\lfloor np \rfloor} {}_n C_k$ とすると, $M \leq 1 + \sqrt{\frac{np}{2\pi(1-p)}} p^{-np} (1-p)^{-n(1-p)}$ が成立すること

を示せ. ただし, n が十分大きいとき, スターリングの公式 $n! \approx \sqrt{2\pi e} n^{n+\frac{1}{2}}$ が成り立つとしてよい. また, $\lfloor x \rfloor$ は, x 以下の最大の整数を表す.

設問 4 上の設問において, M 個の 2 元系列のうち, 1 個は実際に送信された符号語である. 残りの $M-1$ 個の 2 元系列の中にランダム符号化で生成された N 個の符号語のどれかが含まれていると, 復号時に曖昧性が生じて復号誤りとなる. 2^n 個の 2 元系列のうち N 個だけが符号として用いられるので, $M-1$ 個の 2 元系列の中に含まれる符号語の個数の期待値 M_A は $M_A = \frac{N}{2^n} (M-1)$ となる. 上の設問の結果を使って, $N \leq \frac{2^{nC}}{n}$ で

あれば, $M_A \leq \sqrt{\frac{p}{2\pi(1-p)n}}$ となり, $n \rightarrow \infty$ のとき $M_A \rightarrow 0$ となることを示せ. ここで,

C は設問 1 で求めた T の通信路容量である.

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

Consider a memoryless communication binary channel $T = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$. Let us assume that we employ random coding for transmitting N messages in $A = \{a_1, a_2, \dots, a_N\}$ through T , where $x_1 x_2 \dots x_n$ ($x_j \in \{0, 1\}$) of length n is a randomly generated code word for $a_i \in A$ ($i = 1, \dots, N$). At the receiver, we will decode the output of T under the assumption that the number of the erroneous digits is at most np , where $p \leq \frac{1}{2}$, and n is sufficiently large.

[Question 1] Show the channel capacity C of T .

[Question 2] Let a random variable Z denote the number of erroneous digits in a received sequence. Then, the probability of Z is $P\{Z = k\} = {}_n C_{n-k} p^k (1-p)^{n-k}$ for $k = 0, 1, \dots, n$. Prove that the expectation $E(Z)$ of Z is np .

[Question 3] Let $M = \sum_{k=0}^{\lfloor np \rfloor} {}_n C_k$ be the number of the sequences of length n whose Hamming distance is no more than np comparing with the received sequences. M can be considered as the number of possible original messages on an average.

Prove that $M \leq 1 + \sqrt{\frac{np}{2\pi(1-p)}} p^{-np} (1-p)^{-n(1-p)}$ is true. You may assume that n is

sufficiently large and use Starling's formula: $n! \approx \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}$. $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x .

[Question 4] In the above question, if only one of these M binary sequences is correct, decoding error may occur while a code word is contained in the remaining $M-1$ sequences. Suppose only N code words of all 2^n possible messages are used, the expected number of messages in A , which is not the correct original message that could have been changed by noise into the received signal is

$M_A = \frac{N}{2^n} (M-1)$. Now, prove that $M_A \leq \sqrt{\frac{p}{2\pi(1-p)n}}$ is true if $N \leq \frac{2^{nC}}{n}$, and

that $M_A \rightarrow 0$ as $n \rightarrow \infty$. Note that C is the channel capacity obtained in [Question 1].

アナログ画像を $f(x, y)$ とし, 対応するデジタル画像を $g(x_i, y_j)$ ($i, j = 1, 2, \dots, N$) とする.

設問1 アナログ画像 $f(x, y)$ からエッジ点を抽出する方法として, 勾配に基づく方法がある. この方法を説明せよ.

設問2 設問1 で説明した方法をデジタル画像 $g(x_i, y_j)$ に適用してエッジ点を検出する方法を説明せよ.

設問3 アナログ画像 $f(x, y)$ からエッジ点を抽出する方法として, ラプラシアンに基づく方法がある. この方法を説明せよ.

設問4 設問3 で説明した方法をデジタル画像 $g(x_i, y_j)$ に適用してエッジ点を検出する方法を説明せよ.

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

Let $f(x, y)$ and $g(x_i, y_j)$ ($i, j = 1, 2, \dots, N$) be an analog image and its digitized image, respectively.

Question 1 Explain an edge detection method for $f(x, y)$ using gradient.

Question 2 Explain the edge detection method for $g(x_i, y_j)$ corresponding to the method answered in Question 1.

Question 3 Explain an edge detection method for $f(x, y)$ using Laplacian.

Question 4 Explain the edge detection method for $g(x_i, y_j)$ corresponding to the method answered in Question 3.

世界中の様々な種類の犬を自動的に分類することを考える。

設問1. 機械で容易に測定・数値化可能な特徴量を複数挙げよ。このような多数の特徴量に対して、次元を圧縮しながら分類に有用な特徴量に変換する多変量統計解析の代表的な手法の概要を述べよ。

設問2. 上記で得られた特徴量に基づいて、 K 個のクラスタに分類する手続きを記述せよ。

犬と猫を自動的に分類することを考える。

設問3. 機械で容易に測定・数値化可能な特徴量を複数挙げよ。このような多数の特徴量に対して、次元を圧縮しながら分類に有用な特徴量に変換する多変量統計解析の代表的な手法の概要を述べよ。特に、設問1で述べた手法との違いについて説明すること。

設問4. 上記の特徴量によるユークリッド距離に基づいて識別器を構成する際に、犬と猫各々について、テンプレートを1つ用意する場合と複数用意する場合で、識別関数及び識別境界がどのようなになるか議論せよ。

設問5. 上記でユークリッド距離以外の距離を用いる方法について述べよ。

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

Consider a problem of automatic classification of a variety of dogs in the world.

Question 1: List several features that can be easily digitalized by certain sensing devices. Describe a representative dimension reduction method that transforms these digitalized features to ones, which are more suitable for classification, using multivariate statistical analysis.

Question 2: Describe a procedure that clusters dogs into K clusters based on the above feature vector.

Consider a problem of automatic classification between dogs and cats.

Question 3: List several features that can be easily digitalized by certain sensing devices. Describe a representative dimension reduction method that transforms these digitalized features to ones, which are more suitable for classification, using multivariate statistical analysis. Explain the difference from the method described for Question 1.

Question 4: We will design a classifier based on the Euclidean distance of the above feature vector. Discuss what may happen to the discriminant function and the decision boundary when we assign a single template or multiple templates for dogs and cats respectively.

Question 5: Describe a method of using a distance measure other than the Euclidean distance.

図 1 に示す 5×4 のマスで構成された迷路環境を移動するロボット (図中 'o') を考える。ロボットは現在位置がノード(1, 1)であること, ゴール (図中 'G') がノード(5, 4)であることを事前知識として与えられている。ロボットは空のマス (図中 '□') へ一回の移動で上下左右いずれかに 1 マス移動できるが, 図中 '■' となっているマスには移動できない。一回の移動のコストは 1 である。

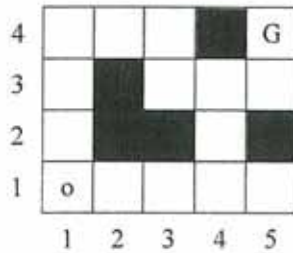


図 1 ロボットが探索する迷路環境

設問 1 この問題について, 幅優先探索, 深さ優先探索を行った場合の探索過程を示しなさい。

設問 2 最適探索と最良探索の違いを説明しなさい。

設問 3 移動中のロボットとゴールの座標から適切(admissible)なヒューリスティック値を設定し, A*探索を用いた探索過程を示しなさい。

設問 4 一般に, ヒューリスティック関数 h が適切(admissible)に設定されているとき, 以下の関数をそれぞれヒューリスティック関数とした場合, 問題が生じるかどうかを理由とともに簡単に述べなさい。

- (1) $2h$
- (2) $h/2$

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

Suppose that a robot is in a maze environment with 5 x 4 cells shown in Figure 1. In advance, the robot ('o' in the figure) knows that its initial position is (1, 1), and the goal ('G' in the figure) is (5, 4). Though the robot can move to an adjacent empty cell ('□' in the figure) with one move, it cannot move to a wall cell ('■' in the figure). The cost of one move is 1.

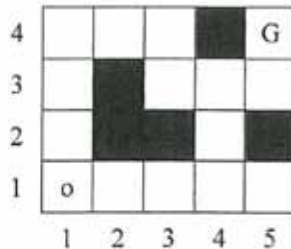


Figure 1: The maze environment the robot needs to explore.

Question 1 Build the search processes when applying depth-first search and breadth-first search to this problem respectively.

Question 2 Explain the difference between optimal search and best-first search.

Question 3 Build a search process when applying A* search to this problem by adding an admissible heuristic value at each cell using the coordinates of the moving robot and goal.

Question 4 Suppose an admissible heuristic h is given. Describe the problems when applying following alternatives.

- (1) $2h$
- (2) $h/2$

n を正整数、 a, b, x を非負整数とする。以下の各設問に答えよ。

設問 1 全加算器(FA)とは 3 個の 1 ビット 2 進数 y, z, w の加算を行い s および桁上げ c を計算する組合せ論理回路である。全加算器の出力 c, s の表す論理関数を乗法標準形 (和積形) で表せ。

設問 2 全加算器と D フリップフロップからなる図 1 のような順序回路を考える。この回路の入力 A および B に、2 個の n ビット符号なし 2 進整数 $a = (a_{n-1}, a_{n-2}, \dots, a_1, a_0)$ および $b = (b_{n-1}, b_{n-2}, \dots, b_1, b_0)$ をビット列として下位桁から逐次的に与えることで、出力 X に 2 数の和 $x = a + b$ の $(n+1)$ ビット 2 進表現 $x = (x_n, x_{n-1}, \dots, x_1, x_0)$ が得られることを説明せよ。ただし初期状態において D フリップフロップはゼロを出力していると仮定する。

設問 3 AND 素子、D フリップフロップおよび図 1 の回路からなる図 2 のような順序回路を考える。この回路の入力 a_3, a_2, a_1, a_0 に 4 ビット符号なし 2 進整数 $a = (a_3, a_2, a_1, a_0)$ の各桁を、また入力 B には 4 ビット符号なし 2 進整数 $b = (b_3, b_2, b_1, b_0)$ をビット列として下位桁から逐次的に与えることで、出力 X に 2 数の積 $x = ab$ の 8 ビット 2 進表現 $x = (x_7, x_6, \dots, x_1, x_0)$ が得られることを説明せよ。ただし初期状態において D フリップフロップはすべてゼロを出力していると仮定する。

設問 4 設問 3 の回路で、初期状態において D フリップフロップの初期値が不定である場合に、設問 3 と同じ計算を行うには入力 B の系列をどのように修正すればよいかを述べよ。またその場合出力はどのように得られるかを述べよ。

設問 5 基本素子として全加算器および AND 素子を用いて、2 個の 4 ビット符号なし 2 進整数 $a = (a_3, a_2, a_1, a_0)$ および $b = (b_3, b_2, b_1, b_0)$ の積 $x = ab$ の 8 ビット 2 進表現 $x = (x_7, x_6, \dots, x_1, x_0)$ を計算する組合せ論理回路を構成せよ。

設問 6 組合せ論理回路の段数とは、入力から出力までの経路上の素子の数の最大値をいう。また組合せ論理回路の素子数とは、回路を構成する素子の総数をいう。設問 5 で構成した回路を任意の n に対する n ビット符号なし乗算回路に一般化する方法について述べ、その段数および素子数を n のオーダ記法 (Big O 記法) で表せ。

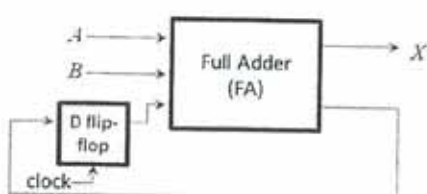


図 1 (Figure 1)

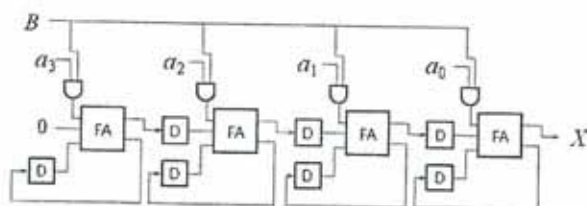


図 2 (Figure 2)

*Question is translated in English in the section below; this translation is given for reference only.
All Figures are given in the previous page.*

Let n be a positive integer, and a , b , and x be non-negative integers. Answer the following questions.

Question 1 A full adder is a combinatorial circuit performing binary addition of three binary digits y , z , w , from which it computes the sum s and the carry-out c . Find each conjunctive normal form (minimal product-of-sum expression) of the Boolean functions which represents the outputs of the full adder c or s respectively.

Question 2 Consider a sequential logic circuit composed of a full adder and a D flip-flop as shown in Figure 1. Two unsigned binary integers with n -bit, $a = (a_{n-1}, a_{n-2}, \dots, a_1, a_0)$ and $b = (b_{n-1}, b_{n-2}, \dots, b_1, b_0)$, are given as inputs of A and B of the circuit, and are operated sequentially from the rightmost bit. The circuit performs addition of the two integers $x = a + b$, and generates an $(n+1)$ -bit binary representation $x = (x_n, x_{n-1}, \dots, x_1, x_0)$ on the output X . Please give an explanation of this operation. Here the D flip-flop is assumed to be holding zero as the output at the initial state.

Question 3 Consider a sequential logic circuit composed of AND gates, D flip-flops and circuits in Figure 1, as shown in Figure 2. Consider that each digit of 4-bit unsigned binary integer $a = (a_3, a_2, a_1, a_0)$ is given as an input of the circuit, a_3, a_2, a_1, a_0 , respectively, and a 4-bit unsigned binary integer $b = (b_3, b_2, b_1, b_0)$ is given as an input of B of the circuit, sequentially from the rightmost bit. The circuit performs multiplication of the two integers $x = ab$ and generates an 8-bit binary representation $x = (x_7, x_6, \dots, x_1, x_0)$ on the output X . Please give an explanation of this operation. Here all D flip-flops are assumed to be holding zero as outputs at the initial state.

Question 4 Consider the same circuit given in Question 3. Assume that the initial states of the D flip-flops are unknown. Explain what modification on the input sequence of B is needed to perform multiplication as the same as the case in Question 3. Also explain how to get the same output in Question 3 in this case.

Question 5 Using full adders and AND gates as elementary gates, design a combinatorial logic circuit which performs multiplication of two 4-bit unsigned integers, $a = (a_3, a_2, a_1, a_0)$ and $b = (b_3, b_2, b_1, b_0)$, and computes the product $x = ab$ in 8-bit binary representation $x = (x_7, x_6, \dots, x_1, x_0)$.

Question 6 The depth of a combinatorial logic circuit is defined as the maximum number of gates on all paths from inputs to outputs of the circuit. The size of a combinatorial logic circuit is defined as the total number of gates of which the circuit is composed. Explain how the circuit designed in Question 5 can be generalized into an n -bit binary multiplier for arbitrary n , and show the depth and the size (the number of gates) of the circuit in Big O (order) notation for n .

確率正規文法 $G = (V, T, P, S)$ がある。ここで S は開始記号である。 V と T はそれぞれ非終端記号と終端記号の有限集合であり、以下のように定義されているとする。

$$V = \{S\}$$

$$T = \{e, g, h, n, o, r, u, w, y, \phi\}$$

さらに、 P は確率が付与された生成規則の有限集合であり、以下の要素からなるとする。

$$S \rightarrow \text{you}S \quad p_1$$

$$S \rightarrow \text{go}S \quad p_2$$

$$S \rightarrow \text{now}S \quad p_3$$

$$S \rightarrow \text{here}S \quad p_4$$

$$S \rightarrow \text{nowhere}S \quad p_5$$

$$S \rightarrow \phi \quad p_6$$

ここで、任意の $1 \leq i \leq 6$ に対して $0 < p_i$ かつ $\sum_{i=1}^6 p_i = 1$ である。例えば、最初の生成規則は、 p_1 の確率で S が $\text{you}S$ に書き換えられることを意味する。

以下の設問に答えよ。

設問 1 文 $\text{yougonowhere}\phi$ のすべての導出木とその確率を示せ。

設問 2 導出木の葉を単語とみなし、ある文の確率最大の導出木をその単語分割結果とすると、前問の文が $\text{you_go_nowhere_}\phi$ のように単語分割される条件を示せ。

設問 3 文法 G が生成するすべての文の生成確率の合計が 1 になることを証明せよ。

Consider a probabilistic regular grammar $G = (V, T, P, S)$, where S is the start symbol. V and T are a set of non terminals and a set of terminals respectively, which are defined as follows:

$$V = \{S\}$$

$$T = \{e, g, h, n, o, r, u, w, y, \phi\}.$$

And P is a set of production rules with a probability, composed of the following elements:

$$\begin{array}{ll} S \rightarrow \text{you}S & p_1 \\ S \rightarrow \text{go}S & p_2 \\ S \rightarrow \text{now}S & p_3 \\ S \rightarrow \text{here}S & p_4 \\ S \rightarrow \text{nowhere}S & p_5 \\ S \rightarrow \phi & p_6, \end{array}$$

where $0 < p_i$ for any $1 \leq i \leq 6$ and $\sum_{i=1}^6 p_i = 1$. For example, the first production rule means that S is rewritten into $\text{you}S$ in the probability p_1 .

Answer the following questions.

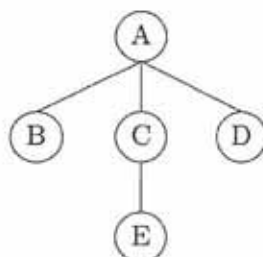
Question 1 Describe all the derivation trees and their corresponding probabilities for the sentence $\text{yougonowhere}\phi$.

Question 2 Considering leaves of the derivation tree of the highest probability of a sentence as words, which are the result of the word segmentation, describe the condition in which the sentence in the previous question is segmented into $\text{you_go_nowhere_}\phi$.

Question 3 Prove that the summation of the probabilities of all the sentences generated by the grammar G is equal to 1.

以下の設問に答えよ。

設問 1 一般の根付き順序木 (rooted ordered tree) を, 二分木 (binary tree) に変換する方法を示せ。ただし, その二分木からもとの根付き順序木が一意に復元できるものとする。また, 下図の木をそのような二分木で表現せよ。



設問 2 二分木のすべての節点 (vertex) を縦型探索 (depth-first search) で走査 (traverse) する方法には下図の例に示す 3 つの方法がある。

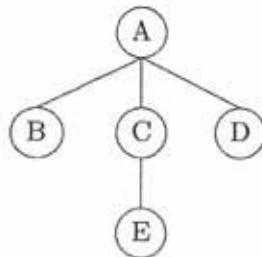


先順, 中順, 後順それぞれの走査方法で節点名を表示する再帰的手続き (recursive procedure) を示せ。ただし, 二分木のデータ構造 (data structure) は適宜定義せよ。

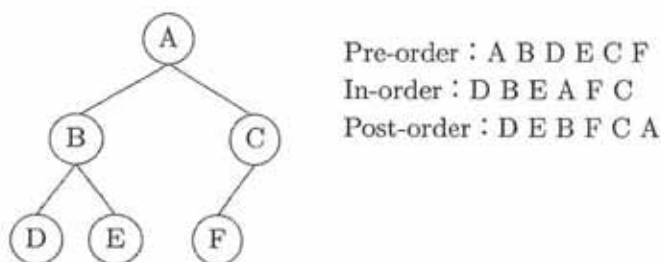
設問 3 二分木の先順と中順の走査結果が与えられれば, 元の二分木が復元できることを示せ。一方, 先順と後順の走査結果が与えられても, 二分木が復元できない具体例を示せ。

Answer the following questions.

Question 1. Explain a method of transforming a rooted ordered tree to a binary tree from which the original rooted ordered tree can be recovered uniquely. Represent the following tree by using such a binary tree notation.



Question 2. As shown in the following example, three methods exist to traverse all the vertices in a binary tree using a depth-first search.



Show recursive procedures to output the labels of vertices by pre-order, in-order and post-order traversals. The data structure of a binary tree also needs to be defined appropriately.

Question 3. Given the pre-order traversal and in-order traversal of a binary tree, explain how the original binary tree can be recovered. On the other hand, show an example that the original binary tree cannot be recovered from the pre-order and post-order traversals.

実変数 t のガウス関数

$$g(t) = e^{-t^2}$$

は、

$$\int_{-\infty}^{\infty} g(t) dt = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

という性質を持つ。以下の設問に答えなさい。

設問 1 正の実数パラメータ σ ($\sigma > 0$) を持つ実変数 t の関数

$$g_{\sigma}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$$

のフーリエ変換を求めなさい。また、 $g_{\sigma}(t)$ の無限積分

$$\int_{-\infty}^{\infty} g_{\sigma}(t) dt$$

の値を求めなさい。

設問 2 $g_{\sigma}(t)$ をフィルタ関数として用い、畳み込み演算によって信号処理を行う場合、フィルタの振幅、位相特性を計算しなさい。ただし、 σ は、フィルタの特性を表わすパラメータとし、 σ の変化によってフィルタの特性がどのように変化するかを述べなさい。

設問 3 畳み込み演算による信号処理では、位相歪が生じることがある。位相歪について説明し、上記の設問2のフィルタにおいて位相歪が生じるか否かについて述べなさい。

設問 4 設問2におけるフィルタ関数 $g_{\sigma}(t)$ をデジタルフィルタとして実現しようとする場合、どのような点について検討する必要があるかを述べなさい。

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

Gaussian function of a real variable t

$$g(t) = e^{-t^2}$$

satisfies the following property.

$$\int_{-\infty}^{\infty} g(t) dt = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

Answer the following questions.

Question 1 Compute the Fourier transform of the following function of the real variable t with a positive real parameter σ ($\sigma > 0$).

$$g_{\sigma}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$$

Compute the following improper integral over the whole real line.

$$\int_{-\infty}^{\infty} g_{\sigma}(t) dt$$

Question 2 Compute the amplitude response and phase response of the convolution-based filtering process using $g_{\sigma}(t)$ as the filter function. Note that σ is the parameter determining the characteristics of the filter. Describe how the filtering characteristics change with respect to changes of σ .

Question 3 Convolution-based filtering processes sometimes introduce phase distortion. Explain what phase distortion implies, and analyze whether or not the filter of Question 2 introduces phase distortion.

Question 4 When we want to implement the filter function of Question 2 $g_{\sigma}(t)$ as a digital filter, describe what technical design problems should be considered.

$G = (V, E)$ を単純無向グラフとする。ただし、 V と E はそれぞれ、 G の頂点集合と枝集合である。

M を E の部分集合とする。 M に含まれるどの2つの枝も同じ頂点を共有しないとき、 M を G の**マッチング**という。マッチング M の**サイズ**とは、 M に含まれる枝の数(すなわち $|M|$)である。 G のマッチングの中で最大サイズのものを G の**最大マッチング**と呼ぶ。 G の最大マッチングのサイズを $X(G)$ と書く。

マッチング M に対して、任意の $e \in E - M$ について $M \cup \{e\}$ がマッチングでないとき、 M は G の**極大マッチング**であるという。 G の極大マッチングの中で最小サイズのものを G の**最小極大マッチング**と呼ぶ。 G の最小極大マッチングのサイズを $Y(G)$ と書く。

枝 $e = (u, v)$ は、頂点 u または v のうち少なくとも一方を端点に持つ枝を**支配**するという(自分自身を支配することに注意)。 D を E の部分集合とする。 E に含まれるどの枝も、 D に含まれるいずれかの枝によって支配されるとき、 D を G の**枝支配集合**と呼ぶ。枝支配集合 D の**サイズ**とは、 D に含まれる枝の数(すなわち $|D|$)である。 G の枝支配集合の中で最小サイズのものを G の**最小枝支配集合**と呼ぶ。 G の最小枝支配集合のサイズを $Z(G)$ と書く。

設問 1. 以下のグラフ G_1 に対して、(i) 最大マッチング、(ii) 最小極大マッチング、(iii) 最小枝支配集合をそれぞれ1つずつ求め、図示せよ。 G_1 の図を解答用紙に描き、答に含まれる枝を太く描くことにより区別せよ。



G_1

設問 2. 8 頂点からなり、最大マッチングのサイズが 1 であるグラフが持ち得る枝の最大数はいくつ? (答のみで良い)

設問 3.

(i) $X(G_2) = 2Y(G_2)$ であるグラフ G_2 を1つ求め、図示せよ。また、理由も示せ。

(ii) 任意のグラフ G に対して $X(G) \leq 2Y(G)$ であることを示せ。

設問 4. 枝支配集合 D に含まれるどの2つの枝も同じ頂点を共有しないとき、 D を**独立枝支配集合**という。以下を命題 1 とする。「 D を G の最小枝支配集合とする。このとき、 $|D'| = |D|$ である G の独立枝支配集合 D' が存在する。」

(i) 命題 1 を使って、任意のグラフ G に対して $Y(G) = Z(G)$ であることを示せ。

(ii) 命題 1 が成り立つことを示せ。

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

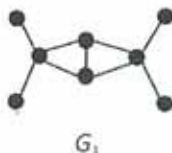
Let $G = (V, E)$ be an undirected simple graph, where V and E are the sets of vertices and edges of G , respectively.

Let M be a subset of E . M is called a **matching** of G if no two edges of M share a common vertex. The **size** of a matching M , denoted by $|M|$, is the number of edges in M . A matching of the largest size is called a **maximum matching** of G . The size of a maximum matching of G is denoted by $X(G)$.

A matching M is called a **maximal matching** if for any edge $e \in E - M$, $M \cup \{e\}$ is not a matching. A maximal matching of minimum size is called a **minimum maximal matching** of G . The size of a minimum maximal matching of G is denoted by $Y(G)$.

An edge $e = (u, v)$ is said to **dominate** any edge containing at least one of u and v (note that any edge dominates itself). Let D be a subset of E . If every edge in E is dominated by at least one edge in D , then D is called an **edge dominating set** of G . The **size** of an edge dominating set D , denoted by $|D|$, is the number of edges in D . An edge dominating set of G with minimum size is called a **minimum edge dominating set** of G . The size of a minimum edge dominating set of G is denoted by $Z(G)$.

Question 1. Find (i) a maximum matching, (ii) a minimum maximal matching, and (iii) a minimum edge dominating set of the following graph G_1 . Give a graphical representation of your answer, i.e., for each of them, draw a graph G_1 using heavy lines for the edges included in your answer.



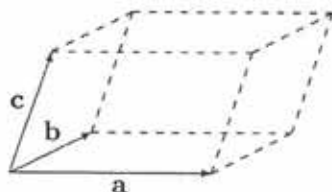
G_1

Question 2. What is the maximum number of edges in a graph that has eight vertices and whose maximum matching size is one? (Give the answer; explanation is not necessary.)

Question 3. (i) Find a graph G_2 such that $X(G_2) = 2Y(G_2)$, and give a graphical representation of G_2 . Briefly explain the correctness of your answer. (ii) Prove that $X(G) \leq 2Y(G)$ for any graph G .

Question 4. An edge dominating set D is called an **independent edge dominating set** if no two edges in D share a common vertex. Let Proposition 1 be the following statement: "For any minimum edge dominating set D of G , there is an independent edge dominating set D' of G such that $|D'| = |D|$." (i) Using Proposition 1, prove that $Y(G) = Z(G)$ for any graph G . (ii) Prove Proposition 1.

設問1 図のように3つのベクトル \mathbf{a} , \mathbf{b} , \mathbf{c} を3辺として作られる3次元空間中の平行六面体の体積に関する以下の問いに答えよ。



- (a) この平行六面体の体積が $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ と表されることを示せ。ただし記号「 $\mathbf{a} \cdot \mathbf{b}$ 」および「 $\mathbf{a} \times \mathbf{b}$ 」はそれぞれ \mathbf{a} と \mathbf{b} の内積, 外積を表すものとする。
- (b) またこの体積が \mathbf{a} , \mathbf{b} , \mathbf{c} を列成分として持つ3行3列の行列 $V = (\mathbf{a} \ \mathbf{b} \ \mathbf{c})$ の行列式 $|V|$ に等しいことを示せ。

設問2 漸化式 $x_{i+1} = x_i^3 - 3x_i^2 + 3x_i$ ($i = 0, 1, 2, \dots$) によって作られる実数列 x_0, x_1, x_2, \dots が収束する (1) 全ての初期値 x_0 の範囲, および (2) それぞれにおける極限值 $\lim_{i \rightarrow \infty} x_i$ を求めよ。

設問3 実関数の導関数および偏導関数について以下の問いに答えよ。

- (a) 微分可能な実関数 $x(t)$, $f(x)$ に対して, 合成関数 $f(x(t))$ の t に関する導関数が

$$\frac{df}{dt} = \frac{df}{dx} \frac{dx}{dt}$$

となることを, 導関数の定義より示せ。

- (b) 実関数 $x(u, v)$, $y(u, v)$, および, 実関数 $f(x, y)$ に対して, $x(u, v) = ve^{u+v}$, $y(u, v) = \frac{v}{u}$, $f(x, y) = xy \log y$ のとき, $f(x, y)$ の u , v それぞれに関する偏導関数を u と v を用いて表せ。

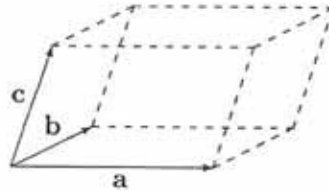
設問4 連続確率分布の一つである指数分布は次の確率密度関数で表される。

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

ただし, $\lambda > 0$ である。確率変数 X が指数分布に従うとき, 以下の問いに答えよ。

- (a) 確率変数 X の累積分布関数 $F_X(x)$ を, その導出過程とともに示せ。
- (b) 確率変数 X の期待値 $E(X)$ を, その導出過程とともに示せ。
- (c) 確率変数 X の分散 $V(X)$ を, その導出過程とともに示せ。

Question 1 Answer the questions concerning the volume of the parallelepiped in \mathbb{R}^3 formed by three vectors \mathbf{a} , \mathbf{b} and \mathbf{c} as illustrated below.



- Prove that the volume is given by $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$. Here $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{a} \times \mathbf{b}$ denote the dot and cross products respectively.
- Prove that the volume is also given by the determinant of the 3×3 matrix $V = (\mathbf{a} \ \mathbf{b} \ \mathbf{c})$. Here the matrix V has \mathbf{a} , \mathbf{b} and \mathbf{c} as the first, second and third column vector respectively.

Question 2 Consider a sequence of real numbers x_0, x_1, x_2, \dots satisfying the recurrence relation $x_{i+1} = x_i^3 - 3x_i^2 + 3x_i$ ($i = 0, 1, 2, \dots$). Find (1) all ranges of x_0 such that each of which makes this sequence converge, and (2) their corresponding $\lim_{i \rightarrow \infty} x_i$.

Question 3 Answer the following questions about derivative and partial derivative of real functions.

- The derivative with respect to t of a composite function $f(x(t))$ of differentiable real functions $x(t)$ and $f(x)$ has the following form.

$$\frac{df}{dt} = \frac{df}{dx} \frac{dx}{dt}$$

Prove this using the definition of the derivative.

- Let real functions $x(u, v)$, $y(u, v)$ and $f(x, y)$ have the following form:
 $x(u, v) = ve^{u+v}$, $y(u, v) = \frac{v}{u}$ and $f(x, y) = xy \log y$. Describe the partial derivative of $f(x, y)$ with respect to each of u and v by using only u and v .

Question 4 An exponential distribution, which is a continuous probability distribution, has the following probability density function.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Here $\lambda > 0$ is the parameter of the distribution. Let a random variable X follow the exponential distribution. Answer the following questions.

- Show the derivation of the cumulative distribution function $F_X(x)$ of the random variable X .
- Show the derivation of the expectation $E(X)$ of the random variable X .
- Show the derivation of the variance $V(X)$ of the random variable X .