

北海道大学 大学院情報科学院

情報科学専攻 修士課程

情報理工学コース

専門科目 2

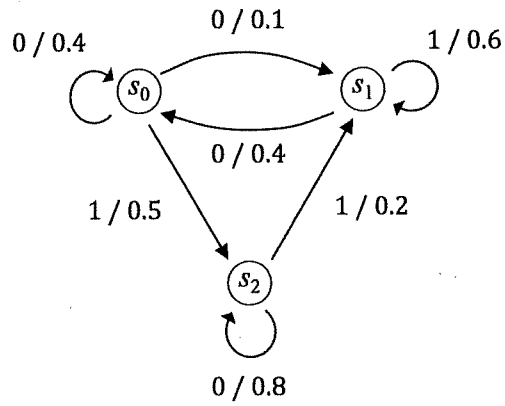
13 : 00 ~ 15 : 00

受験上の注意

- 本冊子には、小問と小論文の計2題ある。それぞれについて解答せよ。ただし、小問については、(1) から (6) の6問のうちから1問を選択して解答すること。
- 小問選択問題チェック票に受験番号および、選択した小問に○印を記入すること。
- 解答用紙は小問用1枚、小論文用1枚の計2枚である。この他に下書き用の草案紙2枚を配付する。
- すべての解答用紙に、受験番号を記入すること。
- 小問用の解答用紙に、選択した小問番号(例えば、(3))を記入すること。
- 解答用紙は裏面を使用してもよい。解答用紙を破損したりした場合には試験監督員に申し出ること。
- 問題冊子、草案紙は持ち帰り、小問選択問題チェック票とすべての解答用紙を提出すること
- 机の上に置いてよいものは、筆記用具(黒鉛筆、消しゴム、鉛筆削り)、時計、および特に指示があったもののみである。時計は計時機能のみを使用し、アラームの使用を禁ずる。携帯電話、スマートフォン、タブレット、コンピュータ等は電源を切ってかばんの中にしまうこと。電卓、電子辞書などは使用を禁ずる。

小問 (1)

次の図で表される 2 元マルコフ情報源 S について、情報源 S のエントロピー $H(S)$ を求めよ。図の各矢印は状態遷移を表し、矢印に付与された 2 つの数字は (出力記号) / (遷移確率) を表す。小数点第 4 位を四捨五入して解答すること。 $\log_2 3 = 1.585$, $\log_2 5 = 2.322$ とする。



小問 (2)

n 個の頂点 (ノード) $V=\{v_1, \dots, v_n\}$ (ただし, $n \geq 2$) と, m 個の辺 (エッジ) $E=\{e_1, \dots, e_m\}$ からなる無向グラフ $G(V, E)$ について, 以下の問いに答えよ. なお, ループ (自己閉路) も多重辺も含まないグラフを単純グラフという.

(1) 以下の定義を答えよ.

(1-1) 頂点の次数

(1-2) 完全グラフ

(2) 単純グラフ $G(V, E)$ のすべての頂点の次数が $\frac{n-1}{2}$ 以上であるとき, $G(V, E)$ は連結であることを示せ.

(3) 辺の数が $m = \frac{(n-1)(n-2)}{2}$ であるような, 連結ではない単純グラフが存在することを示せ.

(4) 単純グラフ $G(V, E)$ の辺の数が $m \geq \frac{(n-1)(n-2)}{2} + 1$ であるとき, $G(V, E)$ は連結であることを示せ.

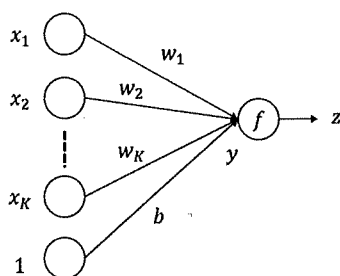
小問 (3)

正整数 a, b ($a \geq b$)を入力とし、それらの最大公約数を出力するユークリッドの互除法によるアルゴリズム $\text{GCD}(a, b)$ について、以下の問いに答えよ。

- (1) 正整数 a, b ($a \geq b$)に対し、 a を b で割った商を q 、余りを r とする。このとき、 b, q, r を用いて a を表せ。
- (2) (1)において $r \neq 0$ のとき、ユークリッドの互除法では $\text{GCD}(a, b)$ の出力が、何と同じであることを用いて計算するのか答えよ。
- (3) アルゴリズム $\text{GCD}(a, b)$ の疑似コードを再帰アルゴリズムで書け。
- (4) (3)の再帰アルゴリズムにおいて、 $\text{GCD}(a, b)$ 実行時の再帰回数が $O(\log_2 a)$ であることを示せ。
- (5) $\text{GCD}(a, b)$ が $\log_2 a$ の多項式時間で計算できることを示せ。ただし、正整数 a, b ($a \geq b$)に対し、 a を b で割った余りは $O((\log_2 a)^2)$ 時間で計算できることを用いること。

小問 (4)

[1] 図で示している K 次元入力で, 1次元出力の中間層のないニューラルネットワークによる回帰問題を考える. サンプル μ において, j 番目の入力を x_j^μ ($j=1, \dots, K$), 出力を z^μ , j 番目の入力と出力の間の結合重みを w_j , 出力のバイアスを b , 出力の目標値を t^μ とする. 出力ノードの入力の総和 y^μ は, 式 (1.1) で計算され, 最終的な出力は, 活性化関数 f を用いて式 (1.2) で計算される. ここでは, 活性化関数は恒等関数 $f(s)=s$ とする.



$$y^\mu = \sum_{j=1}^K w_j x_j^\mu + b \quad (1.1)$$

$$z^\mu = f(y^\mu) \quad (1.2)$$

$$E = \frac{1}{2} \sum_{\mu} (z^\mu - t^\mu)^2 \quad (1.3)$$

誤差関数を (1.3) で示す 2 乗和誤差とし, 勾配法によりバイアス b の最適化を行うために必要となる $\frac{\partial E}{\partial b}$ を導出せよ.

[2] [1] のモデルを使って, 2 クラス分類問題を考える. 入力 x^μ に対する正解ラベルは $t^\mu \in \{0,1\}$ とする. 活性化関数 f を式 (2.1) で示すシグモイド関数にして, 誤差関数 E に式 (2.2) で示す交差エントロピーを用いたときの $\frac{\partial E}{\partial b}$ を導出せよ. ただし, \ln はネイピア数 e を底とする自然対数である.

$$f(s) = \frac{1}{1+e^{-s}} \quad (2.1)$$

$$E = - \sum_{\mu} \left(t^\mu \ln z^\mu + (1-t^\mu) \ln (1-z^\mu) \right) \quad (2.2)$$

[3] 中間層を持つ多層のニューラルネットワークの学習において, 訓練データにはよく適合するが, 訓練データではない未知のデータ (テストデータ) には適合しない過学習の状態に陥ることがある. この過学習を抑制する方法を 2 つ挙げ, その方法について説明せよ.

小問（５）

コンピュータネットワークに関する以下の設問に解答せよ。

- (1) OSI 参照モデルにおけるデータリンク層とネットワーク層の役割の違いを，200 文字程度で説明せよ。なお，説明には MAC アドレスと IP アドレスがそれぞれどの層で利用されるかを含めること。
- (2) DNS はどのような役割をはたすものを 150 文字程度で説明せよ。

小問 (6)

n 次元可微分多様体の定義を述べ, さらに平面上の半径 1 の円周

$$S^1 = \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \}$$

がコンパクト 1 次元可微分多様体となることを定義から確認せよ. さらに工学や情報科学での多様体論の応用事例について知っていることを述べよ.

小論文

近年、インターネットでアクセスできる情報が飛躍的に増えているため、多くの利用者は、検索エンジンの検索結果、ニュースアグリゲーションサイト*1からのおすすめ、興味のある SNS からの情報取得などを中心とした効率の良い情報閲覧を行うようになってきている。一方で、効率を重視するこれらの情報閲覧においては、結果として特定の傾向を持った情報しか見えなくなるといったフィルターバブルと呼ばれる事象が発生し、その社会的な影響が問題となってきている。このようなフィルターバブルの実例とその影響への対策を、次の(a),(b),(c)の3つの項目を含む形で、800文字程度で記述せよ。

- (a) 具体的なサービス（検索エンジン、ニュースアグリゲーションサイト、SNS など）における実例の説明
- (b) (a)の事例が引き起こす社会的な影響についての説明
- (c) (b)で述べたフィルターバブルによる社会的な悪影響を軽減するための対策の提案（ユーザが個人で対応する方法、サービス提供者がシステムを変更することで対応する方法のどちらの方法による提案でも良い）

*1 ニュースアグリゲーションサイト：複数のニュースサイトから提供されるニュースを集約・整理してユーザに提供するサイト