MATH6005 Semester 2, 2024
Assignment 3

There are four questions. You may find some questions more difficult/time-consuming than others, but nevertheless each question is worth the same, 5 marks. Your total score out of 20 marks will determine 10% of your final grade. The assignment is due by 16:00 on the Monday of Teaching Week 12 (21st of October). Submission instructions will be available on Wattle at least one week before the due time. Late submissions are not accepted without an extension. You are strongly advised to submit your assignment well before the due time to allow for unforeseen circumstances.

You should write in complete (English and mathematical) sentences.

Your solutions should be submitted in the form of a .pdf document, with your solution to each new problem starting on a new page. You may typeset your solutions, but you do not need to. More information about how to prepare and submit your solutions will be provided on the class wattle page.

You may discuss your approach to various problems with other students, but you must write your own solutions. For example, you may discuss ideas for how to solve a problem with a classmate, but then you should not keep notes of the discussion, and should write from a blank page to a complete solution on your own. You may NOT use AI in preparing your solutions. It is likely to be unhelpful. To ensure you have followed these instructions, you must include a collaboration statement as part of your assignment submission. Failure to include such a statement may be penalised in the grading scheme. Further instructions will be available on Wattle.

You will submit your assignments through Gradescope. Instructions will be available on Wattle.

Question 1

A Consider a probability experiment in which we roll a six-sided die. The outcome of the experiment is the number facing up on the die after it is rolled. Let $S = \{1, 2, 3, 4, 5, 6\}$ and let $X, Y, Z : S \to \{0, 1, 2, 3\}$ be random variables such that for all $s \in S$ we have

$$X(s) = s \bmod 2, Y(s) = s \bmod 4 \text{ and } Z(s) = \begin{cases} 0 \text{ if } s \leq 4, \\ 1 \text{ otherwise.} \end{cases}$$

I Determine whether or not X and Y are independent. Show your working!

II Determine whether or not X and Z are independent. Show your working!

B A mechanic is working on a car because the front collision detection (FCD) warning light is illuminated and will not turn off. The mechanic looks up the online manual for the make of car and reads

> "When the FCD warning light is illuminated and will not turn off, 90% of the time the cause is a faulty wire, and the other 10% of the time the cause is a faulty front collision sensor."

The online manual suggests that the mechanic should run a diagnostic test on the front collision sensor. The diagnostic test will result in one of two possible messages being displayed on the diagnostic machine: "Fault Found" or "No Fault Found". Regarding the accuracy of the diagnostic test, the online manual says the following:

> "This highly accurate diagnostic test gives the message 'Fault Found' 94% of the time when the front collision sensor is faulty, and just 4% of the time when there is no fault with the front collision sensor."

The mechanic runs the diagnostic and the machine displays the message "Fault Found." What is the probability that the front collision sensor is faulty?

Clearly define your notation when preparing your solution!

Question 2

In a 1983 study[1] of social mobility in England and Wales, the authors collected data on the employment of adult males aged 20-64 years in England and Wales in 1972, and the employment of their fathers. For the study, each type of job a person might hold was classified into one of four sectors, listed here in decreasing order of average salary: (1) Service Sector, (2) Intermediate Sector, (3) Working Class Sector, and (4) Farm Sector. The following data is taken from the study[2]:

	Father's Employment			
Son's Employment	Service	Intermediate	Working Class	Farm
Service	.64	.32	.17	.13
Intermediate	.22	.36	.27	.21
Working Class	.13	.31	.55	.34
Farm	.01	.01	.01	.32

Figure 1: Probabilities of a family moving between sectors in one generation

A Use the data to design a Markov process to model of the social mobility of families that live in England and Wales for many generations. You should clearly state the time steps and the states of the Markov process. You should also write a paragraph discussing any further assumptions you have made so that your model is a Markov process.

B Draw a transition diagram of this Markov process and write down the transition matrix T.

C Calculate T^{10}, T^{100} and T^{101} (or alternatively calculate T^{10}, T^{98} and T^{99}) and make a prediction regarding the steady state of this model.

D Find the steady state vector using the 'shortcut method' taught in lectures. This should match your prediction of the steady state vector from Part (C).

E Use this information to say something about the long-term distribution of jobs taken by adult males from families that remain in England and Wales for many generations.

Hint: For parts (C) and (D) you should use a computer application such as Reshish Matrix Calculator or WolframAlpha, to complete the computations. Include screen shots to show your working.

[1] Anthony Heath and John Ridge, "Social Mobility of Ethnic Minorities", J. Biosoc. Sci., Suppl. 8 (1983) 169-184

[2] The data is taken from Table 4 in the study, and rounded to two decimal places instead of 3.

Question 3

For each $n \in \mathbb{N}$ (positive integers), the hypercube graph H_n is the simple graph formed from the vertices and edges of an n-dimensional hypercube. In particular, the vertex set and edge set can be defined in the following way:

$$V(H_n) = \{v_i \mid i \text{ is a binary string of length } n\}$$
$$E(H_n) = \{\{v_i, v_j\} \mid i \text{ and } j \text{ differ in a exactly one digit}\}$$

Suppose you are designing a parallel computing network consisting of CPUs connected by cables. You propose to arrange the network as a hypercube graph H_n for some $n \geq 1$, where the CPUs correspond to vertices and the cables connecting CPUs correspond to edges.

A Draw diagrams for H_1, H_2, H_3 and H_4. Your diagrams must be neat and vertices should be clearly labeled.

B Your diagnostics consultant is concerned the proposed hypercube arrangement will not allow for the diagnostics team to run diagnostics tests to determine if cable connections are functioning correctly or not. The diagnostics consultant says their testing procedure requires that:

  i. The first cable tested is adjacent to the CPU associated to the binary string of n zeroes.
  ii. Each cable is tested one at a time, in a sequence so that the next cable connects to the same CPU as the previous cable.
  iii. Each cable is tested exactly once.

  Decide whether or not the hypercube arrangement allows the diagnostics tests to be run. Explain how you reached this conclusion!

C Your budget for constructing the parallel computing network is \$700,000. Each cable between two CPUs costs \$100 and each CPU costs \$1000. What is the maximal value of n that the project can afford if you use the proposed hypercube graph H_n arrangement.

  Hint: You will need to calculate the number of CPUs and cables for a given n (show working). If necessary, you may use a computer to complete the calculation.

Question 4

Recall the following:

- A **tree** is defined as a connected graph with no non-trivial circuits.
- A **forest** is a disjoint collection of trees.
- A **leaf** of a tree or forest is a vertex of degree 1.

Your lecturer has told you the following result in lectures:

Lemma 1. *Every tree with more than one vertex has at least one leaf.*

You may use Lemma 1 to answer Part A.

Recall the following theorem introduced in lectures:

Theorem 2 (Tree Characterisations)**.** *Let T be a nonempty graph with n vertices. The following statements are logically equivalent:*

(1) T is a tree.

(2) T has no simple circuits and $n-1$ edges.

(3) T is connected and has $n-1$ edges.

(4) T is connected and every edge is a bridge.

(5) Any two vertices of T are connected by exactly one simple path.

(6) T contains no non-trivial circuits, but the addition of any new edge that connects an existing pair of vertices will create a simple circuit.

We will investigate aspects of proving Theorem 2 and why having such a list of equivalent statements is useful.

A Prove (1) $\implies$ (2).

Hint: Use mathematical induction on the number of vertices ($|V(T)| \geq 1$). This will require you to use Lemma 1.

B Explain why proving each of the statements in the following set results in each of the statements (1)-(6) being logically equivalent to each other and so proves Theorem 2:

$$S = \Big\{(1) \implies (2), (2) \implies (3),\ (3) \implies (4), (4) \implies (5), (5) \implies (6), (6) \implies (1)\Big\}.$$

C You and four other classmates have each proven one of the statements in

$$T = \Big\{(1) \implies (2), (5) \implies (3), (6) \implies (5), (2) \implies (4), (3) \implies (1)\Big\}.$$

Your demonstrator tells you that only proving the statements in T is not sufficient to prove Theorem 2. What $(i, j) \in \{1, 2, 3, 4, 5, 6\}^2$ would make $T \cup \{(i) \implies (j)\}$ a sufficient set of statements to prove Theorem 1. Explain your answer.

D Let G be a graph with s connected components. Prove the statement:

$$\text{If } G \text{ is a forest, then } |E(G)| = |V(G)| - s,$$

Hint: First assume G is a forest and use (2) or (3) to prove $|E(G)| = |V(G)| - s$.

Note: The converse is also true, a proof of which uses (4). This highlights the power of having theorems like Theorem 2 (DO NOT prove the converse in this assignment).