Throughout this exam we write $\mathbb{N}$ for the set $\{1, 2, 3, \ldots\}$ and we write $\mathbb{Z}_{\geq 0}$ for the set $\{0, 1, 2, 3, \ldots\}$.

Problem 1 (10 marks)

(a) For each of the following sentences, circle one of the words STATEMENT/PREDICATE/NEITHER to indicate the nature of the sentence.

(i) $\forall t \in \mathbb{Z} \;\; t \geq t^2$ — STATEMENT/PREDICATE/NEITHER

(ii) $\exists n \in \mathbb{N} \;\; x^2 - \frac{x}{2} = n$ — STATEMENT /PREDICATE/NEITHER

(iii) Google's PageRank algorithm is a clever example of discrete mathematical modelling. STATEMENT/PREDICATE /NEITHER

(iv) For any simple undirected graph G, the total degree of G equals twice the number of edges in G. STATEMENT /PREDICATE/NEITHER

(b) Consider the following statement:

> "Today is not Tuesday nand today is not Tuesday, nand, I am enjoying my exam nand I am enjoying my exam."

Write a statement that is logically equivalent to the statement under consideration, but which only uses the logical structure of an implication. Your statement should be written in English and you should use a truth table to justify that your statement is logically equivalent to the statement under consideration.

Let

$$D: \text{ Today is Tuesday} \qquad E: \text{ I am enjoying my exam}$$

The statement under consideration is written symbolically as follows: $(\neg D \uparrow \neg D) \uparrow (E \uparrow E)$. We compute a truth table:

D	E	$\neg D$	$\neg D \uparrow \neg D$	$\neg D$	$(\neg D \uparrow \neg D) \uparrow (E \uparrow E)$	$E \uparrow E$
T	T	F	T	F	**T**	F
T	F	F	T	F	**F**	T
F	T	T	F	T	**T**	F
F	F	T	F	T	**T**	T

The entries in the truth table agree with those for the statement $D \rightarrow E$. Thus the statement under consideration is logically equivalent to:

> If today is Tuesday, then I am enjoying my exam.

(c) Consider the predicate

$$P(f): \quad \forall a_1 \in \mathbb{N} \ \ \forall a_2 \in \mathbb{N} \ \ \forall b \in \mathbb{N} \ \ (((a_1, b) \in f \wedge (a_2, b) \in f) \Rightarrow (a_1 = a_2)),$$

defined over the domain of functions from $\mathbb{N}$ to $\mathbb{N}$.

(i) Without using the symbol $\neg$, write down a predicate $Q(f)$, defined over the domain of functions from $\mathbb{N}$ to $\mathbb{N}$, such that $Q(f) \equiv \neg P(f)$

$$\begin{aligned} Q(f) &\equiv \neg P(f) \\ &\equiv \neg(\forall a_1 \in \mathbb{N} \ \ \forall a_2 \in \mathbb{N} \ \ \forall b \in \mathbb{N} \ \ (((a_1, b) \in f \wedge (a_2, b) \in f) \Rightarrow (a_1 = a_2))) \\ &\equiv \exists a_1 \in \mathbb{N} \ \ \exists a_2 \in \mathbb{N} \ \ \exists b \in \mathbb{N} \ \ \neg(((a_1, b) \in f \wedge (a_2, b) \in f) \Rightarrow (a_1 = a_2))) \\ &\equiv \quad \exists a_1 \in \mathbb{N} \ \ \exists a_2 \in \mathbb{N} \ \ \exists b \in \mathbb{N} \ \ ((a_1, b) \in f \wedge (a_2, b) \in f \wedge (a_1 \neq a_2)) \end{aligned}$$

(ii) Write down an example of a function $g : \mathbb{N} \to \mathbb{N}$ for which $P(g)$ is true. Justify your answer.

We note that $P(f)$ is logically equivalent to "f is injective."

Let $g : \mathbb{N} \longrightarrow \mathbb{N}$ be given by $g(n) = n$.

Now we prove that g is injective. Let $m, n \in \mathbb{N}$. Suppose that $g(m) = g(n)$. Then $m = g(m) = g(n) = n$. Hence g is injective and $P(g)$ is true.

(iii) Write down an example of a function $h : \mathbb{N} \to \mathbb{N}$ for which $P(h)$ is false. Justify your answer.

Let $h : \mathbb{N} \longrightarrow \mathbb{N}$ be given by $h(n) = 1$.

Since $h(1) = 1 = h(2)$, h is not injective and $P(h)$ is false.

(d) (i) In no more than two sentences, explain what it means to say that a compound statement c is in **disjunctive normal form**.

A compound statement is in disjunctive normal form it is a disjunction of conjunctions, and in each of the conjunctions each statement variable or its negation appears but not both.

(ii) In no more than three sentences, describe an application of the following fact to circuit design:

> Every statement that is not a contradiction is logically equivalent to a statement that is in disjunctive normal form.

Given an input/output table with at least one instance of output 1, the fact given tells us we may write down a compound statement c in disjunctive normal form so that the truth table for c matches the given input/output table (with T's in place of 1's and F's in pace of 0's). Using the compound statement, we can make a circuit based on c and that circuit will produce the given input/output table. Since only $\wedge, \vee, \neg$ are used in disjunctive normal form, our circuit uses only AND, OR and NOT logic gates.

Problem 2 (10 marks)

(a) For each $i \in \mathbb{N}$, we write $M_i = \{\dots, -3i, -2i, -i, 0, i, 2i, 3i, \dots\}$.

(i) Use set-builder notation to describe M_5.

$M_5 = \{z \in \mathbb{Z} \mid z \equiv 0 \pmod 5\}$

ALTERNATIVE ANSWER

$M_5 = \{z \in \mathbb{Z} \mid \exists k \in \mathbb{Z} \;\; z = 5k\}$.

(ii) Either prove or disprove the following statement:

$$\forall n \in \mathbb{N} \;\; (11^n - 6 \in M_5).$$

Let $n \in \mathbb{N}$. We note that $11 \equiv 1 \pmod 5$. We compute

$$\begin{aligned} 11^n - 6 &\equiv 1^n - 6 \pmod 5 && \text{(because } 11 \equiv 1 \pmod 5\text{)} \\ &\equiv 1 - 6 \pmod 5 && \text{(because } 1^n = 1\text{)} \\ &\equiv 1 - 1 \pmod 5 && \text{(because } 6 \equiv 1 \pmod 5\text{)} \\ &\equiv 0 \pmod 5 \end{aligned}$$

Since $11^n - 6 \equiv 0 \pmod 5$. Hence $11^n - 5 \in M_5$.

ALTERNATIVE ANSWER

Let $P(n) : 11^n - 6 \in M_5$. We shall use mathematical induction to prove $\forall n \in \mathbb{N} \;\; P(n)$.

Base case: Since $11^1 - 6 = 11 - 6 = 5 \in M_5$, $P(1)$ holds.

Inductive case: Let $n \in \mathbb{N}$. Suppose that $P(1), P(2), \dots, P(n)$ all hold. Since $P(n)$ holds, $11^n - 6 \in M_5$. Hence there exists $k \in \mathbb{Z}$ such that $11^n - 6 = 5k$. It follows that $11^n = 5k + 6$. Now

$$\begin{aligned} 11^{n+1} - 6 &= 11 \times 11^n - 6 \\ &= 11 \times (5k + 6) - 6 && \text{(Using } P(n)\text{)} \\ &= 55k + 66 - 6 \\ &= 55k + 60 \\ &= 5(11k + 12) \end{aligned}$$

Since $k \in \mathbb{Z}$, $11k + 12 \in \mathbb{Z}$. Hence $11^{n+1} - 6 \in M_5$. Hence $P(n+1)$ holds.

By the Principle of Mathematical Induction, $\forall n \in \mathbb{N} \;\; P(n)$.

(b) Let E denote the set of even integers.

(i) Let A and B be sets. Write down what it means to say that A and B have the same cardinality.

We say that A and B have the same cardinality when there exists a bijection from A to B.

(ii) Write down the rule for a bijection $f : \mathbb{N} \to E$, or explain how you know that no such bijection exists.

(There are many possible answers. One is...)

$$f(n) = \begin{cases} n \text{ if } n \text{ is even,} \\ 1 - n \text{ if } n \text{ is odd.} \end{cases}$$

(iii) What fact about E does your answer to (ii) demonstrate?

The existence of a bijection from $\mathbb{N}$ to $\mathbb{E}$ demonstrates that E is countably infinite (or just countable).

(c) Either use the logical equivalences below and the definitions of set operations to prove, or provide a counterexample to disprove, the following statement:

For any universal set U, and any sets $A, B, C, \in \mathcal{P}(U)$, $A \cup (B \setminus C) = B \cup (A \setminus C)$.

The following example demonstrates that the statement is false. Let

$$U = \{1, 2\},$$
$$A = \{1\}$$
$$, B = \{2\}, \text{ and}$$
$$C = \{1, 2\}.$$

Then

$$A\cup(B\setminus C) = \{1\}\cup(\{2\}\setminus\{1,2\}) = \{1\}\cup\emptyset = \{1\} \neq \{2\} = \{2\}\cup\emptyset == \{2\}\cup(\{1\}\setminus\{1,2\}) = B\cup(A\setminus C).$$

(To see that the statement is false, and to see how to make a counterexample, make two Venn diagrams and note the differences between them.)

Given any statement variables p, q, and r, a tautology t and a contradiction c, the following hold.

1. Commutative laws:	$p \wedge q \equiv q \wedge p$	$p \vee q \equiv q \vee p$
2. Associative laws:	$(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$	$(p \vee q) \vee r \equiv p \vee (q \vee r)$
3. Distributive laws:	$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$
4. Identity laws:	$p \wedge t \equiv p$	$p \vee c \equiv p$
5. Negation laws:	$p \vee \neg p \equiv t$	$p \wedge \neg p \equiv c$
6. Double negative law:	$\neg(\neg p) \equiv p$	
7. Idempotent laws:	$p \wedge p \equiv p$	$p \vee p \equiv p$
8. Universal bound laws:	$p \vee t \equiv t$	$p \wedge c \equiv c$
9. De Morgan's laws:	$\neg(p \wedge q) \equiv \neg p \vee \neg q$	$\neg(p \vee q) \equiv \neg p \wedge \neg q$
10. Absorption laws:	$p \vee (p \wedge q) \equiv p$	$p \wedge (p \vee q) \equiv p$
11. Negations of t and c:	$\neg t \equiv c$	$\neg c \equiv t$

(d) Throughout this part of the problem we consider the following statement:

For any $x \in \mathbb{Z}$, if $x \not\equiv 2 \pmod 3$ or $x \not\equiv 2 \pmod 5$, then $x \not\equiv 2 \pmod{75}$.

(i) Write down the logical structure of a proof of the statement under consideration that proceeds directly.

Let $x \in \mathbb{Z}$
Suppose that $x \not\equiv 2 \pmod 3$ or $x \not\equiv 2 \pmod 5$.
$\vdots$
Hence $x \not\equiv 2 \pmod{75}$.

(ii) Write down the logical structure of a proof of the statement under consideration that proceeds via the contrapositive.

Let $x \in \mathbb{Z}$
Suppose that $x \equiv 2 \pmod{75}$.
$\vdots$
Hence $x \equiv 2 \pmod 3$ and $x \equiv 2 \pmod 5$.

(iii) Prove or disprove the statement under consideration.

Let $x \in \mathbb{Z}$ Suppose that $x \equiv 2 \pmod{75}$. By definition, there exists $k \in \mathbb{Z}$ such that $x - 2 = 75k$.
Now $x - 2 = 75k = 3(25k)$. Since $k \in \mathbb{Z}$, $25k \in \mathbb{Z}$. Hence $x \equiv 2 \pmod 3$.
Now $x - 2 = 75k = 5(15k)$. Since $k \in \mathbb{Z}$, $15k \in \mathbb{Z}$. Hence $x \equiv 2 \pmod 5$.
Hence $x \equiv 2 \pmod 3$ and $x \equiv 2 \pmod 5$.

ALTERNATIVE ANSWER

Let $x \in \mathbb{Z}$. Suppose that $x \equiv 2 \pmod{75}$. By definition, there exists $k \in \mathbb{Z}$ such that $x - 2 = 75k$. So $x = 75k + 2$.
We compute

$$\begin{aligned} x = 75k + 2 &\equiv 0k + 2 \pmod 3 \qquad &&(\text{because } 75 \equiv 0 \pmod 3)\\ &\equiv 2 \pmod 3 \end{aligned}$$

and

$$\begin{aligned} x = 75k + 2 &\equiv 0k + 2 \pmod 5 \qquad &&(\text{because } 75 \equiv 0 \pmod 5)\\ &\equiv 2 \pmod 5 \end{aligned}$$

Hence $x \equiv 2 \pmod 3$ and $x \equiv 2 \pmod 5$.

Problem 3 (10 marks)

(a) A web-banking password is always 8 characters long and it always comprises two upper-case letters from the standard English alphabet, two digits, and four lower-case letters from the standard English alphabet. How many different passwords can be created that follow these rules?
Give your answer in the form of an expression that uses only arithmetic operations (possibly including factorial and exponentiation) and numbers.

We construct a password using a multistage process.

Step 1 We choose which 2 of 8 positions will be occupied by upper-case letters. There are $\binom{8}{2}$ different possible outcomes from this step.

Step 2 We choose which 2 of the remaining 6 positions will be occupied by digits. There are $\binom{6}{2}$ different possible outcomes from this step.

Step 3 We choose the upper-case letter to go in the first (left-most) position to be occupied by an upper-case letter. There are 26 different possible outcomes from this step.

Step 4 We choose the upper-case letter to go in the second position to be occupied by an upper-case letter. There are 26 different possible outcomes from this step.

Step 5 We choose the lower-case letter to go in the first (left-most) position to be occupied by a lower-case letter. There are 26 different possible outcomes from this step

Step 6 We choose the lower-case letter to go in the second position to be occupied by a lower-case letter. There are 26 different possible outcomes from this step.

Step 7 We choose the lower-case letter to go in the third position to be occupied by a lower-case letter. There are 26 different possible outcomes from this step.

Step 8 We choose the lower-case letter to go in the fourth position to be occupied by a lower-case letter. There are 26 different possible outcomes from this step.

Step 9 We choose the digit to go in the first (left-most) position to be occupied by a digit. There are 10 different possible outcomes from this step.

Step 10 We choose the digit to go in the second position to be occupied by a digit. There are 10 different possible outcomes from this step.

By the product rule, the total number of different outcomes from this process is

$$\binom{8}{2} \times \binom{6}{2} \times 26^6 \times 10^2 = \frac{8!}{6!2!} \times \frac{6!}{4!2!} \times 26^6 \times 10^2.$$

(b) 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 $X, Y : S \longrightarrow \{0, 1\}$ be random variables such that for all $s \in S$ we have

$$X(s) = s \bmod 2, \text{ and } Y(s) = \begin{cases} 0 \text{ if } s \leq 3, \\ 1 \text{ otherwise.} \end{cases}$$

Determine whether or not X, Y are independent. Justify your answer.

We assume the die is fair, and so the outcomes are equally likely. Since the sample space is finite and outcomes are equally likely, for any event $E \subset S$ we have $\mathbb{P}(E) = \dfrac{|E|}{|S|}$.

Recall that X, Y are independent if and only if

$$\forall x \in \text{Range}(X) \;\; \forall y \in \text{Range}(Y) \;\; \mathbb{P}(X(s) = x \text{ and } Y(s) = y) = \mathbb{P}(X(s) = x) \times \mathbb{P}(Y(s) = y).$$

We compute

$$\begin{aligned} \mathbb{P}(X(s) = 0 \text{ and } Y(s) = 0) &= \frac{|\{2\}|}{|\{1, 2, 3, 4, 5, 6\}|} = \frac{1}{6} \\ \mathbb{P}(X(s) = 0) &= \frac{|\{2, 4, 6\}|}{|\{1, 2, 3, 4, 5, 6\}|} = \frac{3}{6} = \frac{1}{2} \\ \mathbb{P}(Y(s) = 0) &= \frac{|\{1, 2, 3\}|}{|\{1, 2, 3, 4, 5, 6\}|} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

Since $\dfrac{1}{6} \neq \dfrac{1}{4} = \dfrac{1}{2} \times \dfrac{1}{2}$, we have that X, Y are dependent.

(c) A 4-tuple $(x_1, x_2, x_3, x_4) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ is **valid** if $x_1 + x_2 + x_3 + x_4 = 50$. If a valid 4-tuple is chosen at random, what is the probability that it will contain the numbers 5, 10, 15 and 20 in some order?

Give your answer in the form of an expression that uses only arithmetic operations (possibly including factorial and exponentiation) and numbers.

In an excellent response, each of the following will be clearly identified: the experiment, the sample space, how probabilities will be computed, and any events of interest.

The experiment is to select a tuple at random from the set of all valid 4-tuples.
The outcome is the 4-tuple chosen.
The sample space is the set S of all valid 4-tuples.
Since the tuple is chosen at random, each outcome is equally likely.
Since the sample space is finite and outcomes are equally likely, for any event $E \subset S$ we have $\mathbb{P}(E) = \dfrac{|E|}{|S|}$.
Let E denote the event that the 4-tuple chosen contains the numbers 5, 10, 15 and 20 in some order.

Each valid 4-tuple may be represented uniquely by an arrangement of 3 bars and 50 stars. The bars separate the line into 4 compartments, and the number of stars in the respective compartment indicates the value of x_1, x_2, x_3 and x_4. Thus we have

$$|S| = \#\text{ ways to arrange 3 bars and 50 stars in a line } = \binom{53}{3} = \frac{53!}{50!3!} = \frac{53 \times 52 \times 51}{6}.$$

The number of tuples in E is equal to the number of ways to arrange the numbers 5, 10, 15, 20 in order. Thus we have

$$|E| = 4! = 24$$

and

$$\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{24}{\left(\dfrac{53 \times 52 \times 51}{6}\right)} = \frac{24 \times 6}{53 \times 52 \times 51} = \frac{144}{53 \times 52 \times 51}.$$

Problem 4 (10 marks)

(a) Your friend tries to define the term “graph” and writes the following:

> “A graph G is an ordered pair $G = (V(G), E(G))$ comprising: a set of vertices $V(G)$; and a set of edges $E(G)$, with each edge being a set of 2 vertices.”

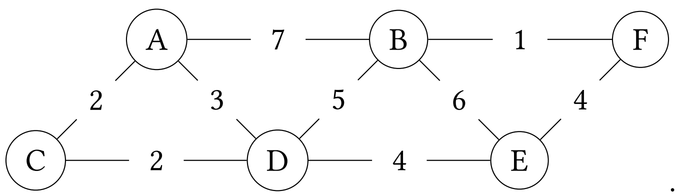
Looking at our lecture notes, you see that we defined the term “graph” as follows:

> “A graph G is an ordered pair $G = (V(G), E(G))$ comprising: a set of vertices $V(G)$; and a multiset of edges $E(G)$, with each edge being a size-2 multiset of vertices.”

In no more than three sentences and using appropriate examples, identify the effective difference(s) between the two definitions.

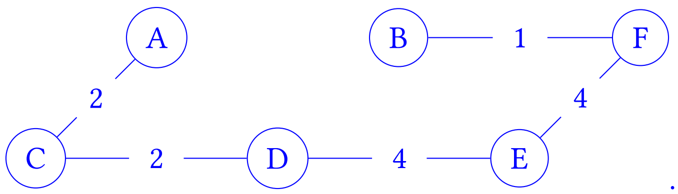
Since a multiset allows repetition while a set does not, the lecture definition allows edges to be repeated and for both endpoints of an edge to be the same while my friend’s definition does not allow these things. The effect is that the lecture definition allows parallel edges and loops, while my friend’s definition allows neither (in fact, my friends definition only allows ‘simple graphs’ to be ‘graphs’). For example, using the lecture definition we may define a graph G with vertex set $\{a, b\}$ and a multiset of edges $E(G) = \{\{a, a\}, \{a, b\}, \{a, b\}\}$; this would be not be a graph under my friend’s definition.

(b) Let G be the weighted graph



Draw a minimal spanning tree for G in the space below.

Using Kruskal’s algorithm, or by inspection, we obtain:



(c) Let

$$E = \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid K_{m,\,n} \text{ has an Euler circuit and } m + n \geq 3\}$$
$$H = \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid K_{m,\,n} \text{ has a Hamilton circuit and } m + n \geq 3\}.$$

(i) Describe H in set-roster notation. Prove that your description gives H.

$H = \{(2, 2), (3, 3), (4, 4), \ldots\}$

Proof: Let $m, n \in \mathbb{N}$ be such that $m+n \geq 3$ (this excludes only the pair $(1, 1)$. The vertices of $K_{m,n}$ may be partitioned into sets $A = \{a_1, a_2, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_n\}$ so that every set of the form $\{a_i, b_j\}$ is an edge and there are no other edges.

Suppose that $m = n$. It is easily checked that the following is a Hamilton circuit in $K_{m,n}$:

$$a_1 b_1 a_2 b_2 \ldots a_m b_m a_m.$$

Hence $K_{m,n}$ has a Hamilton circuit.

Now suppose that $K_{m,n}$ has a Hamilton circuit. Since the only edges in $K_{m,n}$ are of the form $\{a_i, b_j\}$, the circuit must alternate vertices from A and B. In order to alternate vertices from A and B while writing each vertex, A and B must have the same cardinality; that is, $m = n$.

(ii) Use set-roster notation to describe $E \cap H$. Justify your answer.

Recall that a connected graph has an Euler circuit if and only if every vertex has even degree. In the notation of our solution to (i), each vertex in A has degree n and each vertex in B has degree m. It follows that

$$E = \{(m, n) \in \mathbb{N} \times \mathbb{N} \mid m \text{ and } n \text{ are both even and } m + n \geq 3\}.$$

Hence $E \cap H = \{(2, 2), (4, 4), (6, 6), \ldots\}$.

(d) In this problem we consider the vertex-labelling algorithm described in the course (pseudocode for the algorithm is given on page 21).

(i) In no more than three sentences, describe a physical system that may be reasonably modelled by a transport network. In your description, clearly state what it is about the physical system that the vertices and directed edges and capacities represent.

(There are many possible solutions. Here is one...) We may model a sewer system for a city as a transport network. Each vertex represents a junction or a reservoir; each directed edge represents a pipe through which fluid may flow; and the capacity of an edge represents the rate at which fluid may flow through the pipe.

(ii) Is the vertex-labelling algorithm an example of a greedy algorithm? Justify your answer.

The vertex-labelling algorithm is not a greedy algorithm. When assigning levels and labels we may our choices based on the alphabetically first possible flows we find. This means that we do not construct the best (highest flow) incremental flow possible at each stage.

(iii) The following is a quote from our lecture slides:

> "Given a relation $R \subseteq S \times T$, a **matching problem** seeks an injective function $f : S' \to T$ with domain $S' \subseteq S$ as large as possible subject to m being an injective subset of R."

In no more than three sentences explain how a matching problem for $R \subseteq S \times T$ may be converted to the maximum flow problem for a transport network (so that the vertex-labelling algorithm may be used to solve the matching problem).

We construct a transport network as follows:

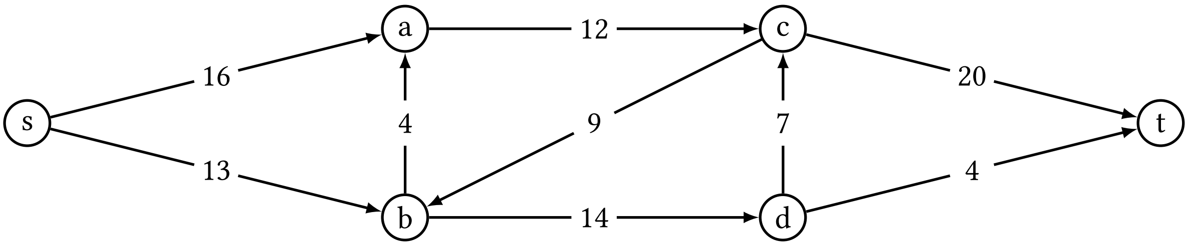
- We start with a vertex for each element of S and a vertex for each element of T, together with a directed edge (s, t) when $(s, t) \in R$;
- We add a super-source s';
- We add a new directed edge (s', s) for each $s \in S$;
- We add a super-target t';
- We add a new directed edge (t, t') for each $t \in T'$;
- We assign each directed edge capacity 1.

Given maximum flow across this transport network, the set

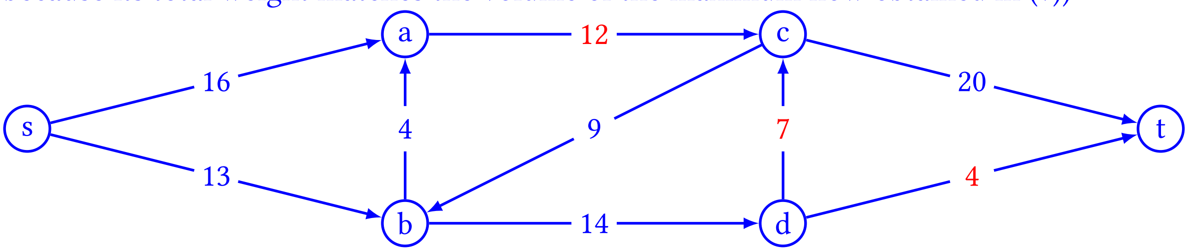
$$\{(s, t) \in S \times T | \text{there is flow across the edge } (s, t)\}.$$

is a solution to the matching problem.

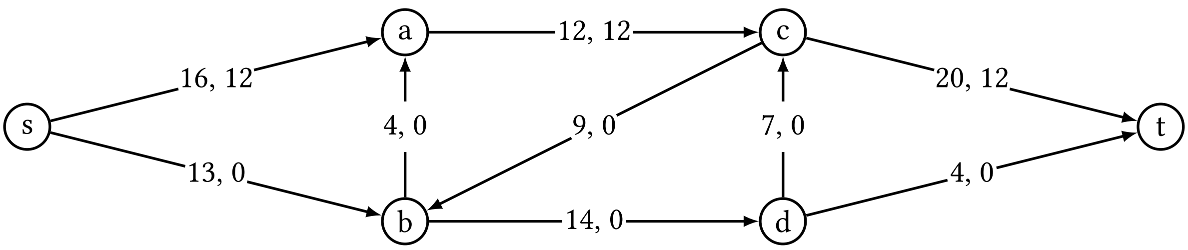
(iv) Indicate a minimum cut on the following transport network.



The cut comprises the edges with labels written in red. (We know it is a minimum cut because its total weight matches the volume of the maximum flow obtained in (v)):



(v) Use the vertex-labelling algorithm described in the course to find a maximum flow function for the transport network in part (iv). The first incremental flow f_1 is shown in the first row of the table at the bottom of the page, and the cumulative flow F_1 is shown in the graph below. Write down the subsequent incremental flows in the table (use only as many rows as you need).

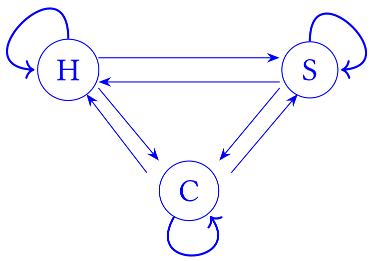


incremental flow label	path of incremental flow	volume of incremental flow
f_1	$s\ a\ c\ t$	12
f_2	$s\ b\ d\ t$	4
f_3	$s\ b\ d\ c\ t$	7

Problem 5 (10 marks)

(a) We consider population movement within a geographic region comprising a city, its suburbs, and the hinterland. We aim to understand the relative populations of the city, the suburbs, and the hinterland as they will change over time. Suppose that the annual migration between the three parts of the geographic region obeys the following. Each year: 4% of the city residents move to the suburbs and 2% move to the hinterland; 7% of the suburban residents move to the city and 2% move to the hinterland; 4% of hinterland residents move to the city and 6% move to the suburbs; all other residents stay in the region they were in.

(i) Draw a transition diagram for a Markov process model of the population movement.



With appropriate weights attached!

(ii) Write down a transition matrix T for your model.

With states in the order: city, suburbs, hinterland we have

$$T = \begin{bmatrix} 0.94 & 0.07 & 0.04 \\ 0.04 & 0.91 & 0.06 \\ 0.02 & 0.02 & 0.90 \end{bmatrix}$$

(iii) List all of the properties a 3×1 vector $\mathbf{v}$ must satisfy in order to be a steady state of your model?

We must have that $\mathbf{v}$ is a probability vector (its entries are non-negative and sum to 1), and $T\mathbf{v} = \mathbf{v}$.

(b) In this problem we consider applying Google's PageRank algorithm to a small webgraph W. The adjacency matrix for W is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

(i) We note that A is a 7×7 matrix. What does this tell us about the web represented by W?

The fact that A is a 7×7 matrix tells us that the web has 7 pages.

(ii) We note that the entry in row 2 column 4 of the matrix A is 1. What does this tell us about the web represented by W?

The fact that the entry in row 2 column 4 of the matrix A is 1 tells us on page 2 there is a hyperlink to page 4.

(iii) Write down the basic transition matrix T when Google's PageRank algorithm is applied to W.

$$T = \begin{bmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{4} & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{4} & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{4} & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{6} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{6} & 0 & 0 & 0 \end{bmatrix}.$$

(iv) Write down an expression for the modified transition matrix M when Google's PageRank algorithm is applied to W. Use a damping factor of 85% (that is, $\alpha = 0.15$).

$$M = \frac{0.15}{7}U + 0.85T,$$

where U is the 7×7 matrix with each entry 1.

(v) In no more than three sentences, explain how to use the matrix M to rank the pages in our small web in order of "awesomeness" (or importance).

We find the unique probability vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_7 \end{bmatrix}$ that is a steady state of the Markov process corresponding to M (that is, we solve $M\mathbf{v} = \mathbf{v}$ subject to $v_1 + v_2 + \cdots + v_7 = 1$). Pages are then ranked in non-increasing order of the corresponding v_i's; that is, the bigger v_i is, the more awesome page i is.

Vertex-labelling algorithm for finding a maximum flow function for a transport network

Input: Transport network D with capacity function C.
Output: A maximum flow function $F_{\max}$ for the network.
Method: Initialise F to the zero flow F_0. Initialize i to 1.

For $i = 1, 2, \ldots$ carry out stage i below to attempt to build an incremental flow f_i.

If stage i succeeds, define $F_i = F_{i-1} + f_i$ and proceed to stage $i+1$.

If stage i fails, define $F_{\max} = F_{i-1}$ and stop.

Stage i:

(a) If $i > 1$, mark up the amended edge flows for F_{i-1}.

(b) Mark up the levels for F_{i-1}, as explained below.

(c) If t is assigned a level, stage i will succeed, so continue.

If not, then stage i fails, so return above to define $F_{\max}$ and terminate.

(d) Mark up labels for F_{i-1} as follows until t is labelled:

(i) Label each level 1 vertex v with sk_{v}, where $k_{\text{v}} = S((s,v))$. (see below for definition of S)

(ii) If t has level 2 or more now work through the level 2 vertices in alphabetical order, labelling each vertex v with uk_u, where

- u is the alphabetically earliest level 1 vertex with $(u,v) \in E(D)$ and $S((u,v)) > 0$,
- k_{u} is the minimum of $S((u,v))$ and the value part of u's label.

(iii) If t has level 3 or more now work through the level 3 vertices in a similar manner and so on.

(e) Let p_i be the path $\text{u}_0\text{u}_1 \ldots \text{u}_n$ where u_n = t and for $0 < j \leq n$
u_j has label $\text{u}_{j-1}k_j$.

Define f_i to be the incremental flow on p_i with flow value k_n.

End of Method
Levels and labels: At each stage of the vertex labelling algorithm levels and labels are associated afresh with the vertices of the network.

The **level** of a vertex is determined iteratively as follows:

- The source vertex s always has level 0.
- If $e = (s, x)$ and $S(e) > 0$ then x has level 1.
- If x has level n and $S((x, y)) > 0$ then y has level $n + 1$ provided it has not already been assigned a lower level.

The **label** on a vertex v of level $n \geq 1$ has the form uk, where u is a vertex of level $n - 1$ and $(u, v) \in E(D)$ is an edge on the path for a potential incremental flow through v with flow value k.
The algorithm assigns labels in ascending order of levels, and in alphabetical order within levels.

The spare capacity function S
For vertices u,v of D, where D has capacity and flow functions C, F:

$$S((\text{u,v})) = \begin{cases} C((\text{u,v}))-F((\text{u,v})) & \text{if (u,v)} \in E(D) \\ F((\text{v,u})) & \text{if (v,u)} \in E(D) \\ 0 & \text{otherwise.} \end{cases}$$

When (v,u) $\in E(D)$, $S((\text{u, v}))$ is called a **virtual capacity**.

This page is intentionally blank (apart from the awesome custom art below, drawn by your classmate).
If necessary, you may use this space for extra working.