数学workshop

Week 1

Question 2 Negate each of the statements below.
Use as natural sounding English as you can manage, and try to avoid using the word 'not'. Do not use symbols.

(a) She will win silver or gold.

(b) She will win silver if she fails to win gold.

(c) She will win gold in the 100m event and in the 200m event.

(a) Directly: She will not win silver and not win gold.

One of several more natural ways to say this:

She will win neither silver nor gold.

(b) The given statement has the form $\neg g \implies s$*, which is equivalent to* $\neg\neg g \vee s$ *and hence to* $s \vee g$*, which is the statemnt given for (a). So the negation is*

Same answer as for (a).

(c) Directly: She will not win gold in the 100m event or she will not win gold in the 200m event.

NB: *The statement "She will not win gold in the 100m event or in the 200m event" is at best ambiguous and in practice usually interpreted as a conjunction ("She won't win gold in either event"), which is* ***not*** *the negation of the given statement.*

One of several more succinct and precise negations is:

She will win at most one gold from the 100m and 200m events.

Question 4 Let p : "If the new drug succeeds, diabetes rates will fall".

(a) Write out the converse of p. Is this equivalent to p?

(b) Write out the contrapositive of p. Is this equivalent to p?

(c) Express p using the phrase "necessary condition".

(a) (i) Directly: If diabetes rates will fall, (then) the new drug succeeds.

This sounds unnatural English. Better is

$\boxed{\textit{If dabetes rates fall, the new drug will have succeeded.}}$

(ii) This is $\boxed{\textit{not}}$ *equivalent to the original because* $(a \Rightarrow b) \not\equiv (b \Rightarrow a)$.

(b) (i) Negating each component of the answer to (a) above gives

$\boxed{\textit{If diabetes rates don't fall, the new drug will have failed.}}$

(ii) This $\boxed{\textit{is}}$ *equivalent to the original because* $(a \Rightarrow b) \equiv (\neg b \Rightarrow \neg a)$.

(c) For any conditional statement, say $s \Rightarrow n$, *we say that* s *is a sufficient condition for* n *and that* n *is a necessary condition for* s.

In particular, another way to express p *is:*

$\boxed{\textit{Falling diabetes rates is a necessary condition for the drug's success.}}$

[In this context 'requirement' might be a more appropriate word than 'condition'.]

Question 5 For each of the following sentences, say whether the sentence is a true statement, a false statement or a predicate. Also give the negation of each sentence.

(a) If $x^2 > 0$ then $x > 0$.

(b) $\forall x \in \mathbb{N}\ \exists y \in \mathbb{N}\ \ x = y^2$.

(c) $\exists! x \in \mathbb{N}\ \ 3x - x^2 = 2$. [*This one is tricky!*]

(a) *(i) Strictly speaking, this is a* $\boxed{\textit{predicate}}$ *because variable x is unquantified (and so requires a value to be specified before truth can be determined.)*
However, in practice, the context of a sentence like this often implies a tacit *universal quantifier whose domain is to be 'understood' from the context. In this case the resulting statement would be true for domain $\mathbb{N}^\star = \{0, 1, 2, 3, \ldots\}$ but false for domain $\mathbb{R}$.*

(ii) Negation: Strictly speaking, $\boxed{x^2 > 0 \land x \leq 0}$.
But if a tacit quantifier $\forall x$ applies, then the negation needs the quantifier $\exists x$.

(b) *(i) This is a* $\boxed{\textit{false statement}}$. *A counterexample is $x = 2$ because 2 is not the square of any natural number.*

(ii) Negation: $\boxed{\exists x \in \mathbb{N}\ \forall y \in \mathbb{N}\ x \neq y^2}$.
Note that this just asserts the existence of a counterexample.

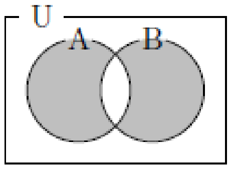
(c) *(i) This statement asserts that the equation $3x - x^2 = 2$ has a unique solution in $\mathbb{N}$. This is a* $\boxed{\textit{false statement}}$ *because in fact the equation has two solutions in $\mathbb{N}$,* viz *1 and 2. (Solve the quadratic equation $x^2 - 3x + 2 = 0$.)*

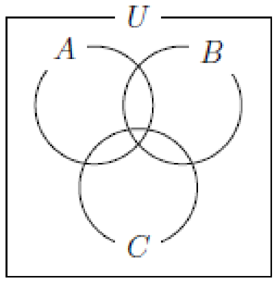
(ii) The negation of unique existence is non-existence or multiple existence. Thus:
$\boxed{(\forall x \in \mathbb{N}\ 3x - x^2 \neq 2) \lor (\exists x, y \in \mathbb{N}\ (3x - x^2 = 3y - y^2 = 2) \land (x \neq y))}$.

Week 2

Question 3 A **Venn diagram** is a graphical reprsentation of sets, subsets and elements. Inside a rectangle for the universal set are one or more circles (or shapes) representing sets. Dots inside the circles (or shapes) represent elements. A subset of the universe, and its relationship to the sets represented by circles, can be indicated by shading various regions on the Venn diagram.

For example, this Venn diagram represents the symmetric difference of A and B.





(a) Using copies of the skeleton Venn diagram at right, draw five diagrams, one for each of the following:

i) $A \cup B$ ii) $(A \cup B)\backslash C$

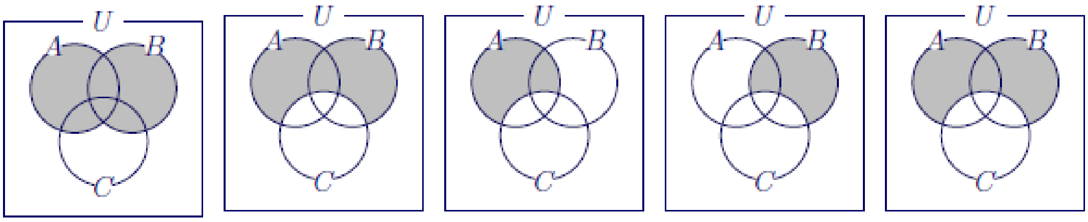
iii) $A\backslash C$ iv) $B\backslash C$ v) $(A\backslash C) \cup (B\backslash C)$

(b) Based on your answers to (a) decide whether $(A \cup B)\backslash C = (A\backslash C) \cup (B\backslash C)$.

(c) Use an element proof (and a logical equivalence) to prove your answer to (b).

(i) *(ii)* *(iii)* *(iv)* *(v)*

(a)



$A \cup B$ $(A \cup B)\backslash C$ $A\backslash C$ $B\backslash C$ $(A\backslash C) \cup (B\backslash C)$

(b) The second and fifth diagrams are the same, so the corresponding sets should be the same. So we decide that $\boxed{\text{the equation is correct}}$.

(c) We prove that $\forall x \in U\ (x \in (A \cup B) \setminus C) \Leftrightarrow (x \in (A \setminus C) \cup (B \setminus C))$. *Let* $x \in U$. *Then*

$x \in (A \cup B)\backslash C$

$\Leftrightarrow x \in (A \cup B)\ \wedge\ \neg(x \in C)$ *(def. of set diff.)*

$\Leftrightarrow [(x \in A) \vee (x \in B)]\ \wedge\ \neg(x \in C)$ *(def. of union)*

$\Leftrightarrow [(x \in A) \wedge \neg(x \in C)] \vee [(x \in B) \wedge \neg(x \in C)]$ ***(logical equivalence)***

$\Leftrightarrow (x \in A\backslash C) \vee (x \in B\backslash C)$ *(def. of set diff.)*

$x \in (A\backslash C) \cup (B\backslash C)$ *(def. of union).*

Question 4

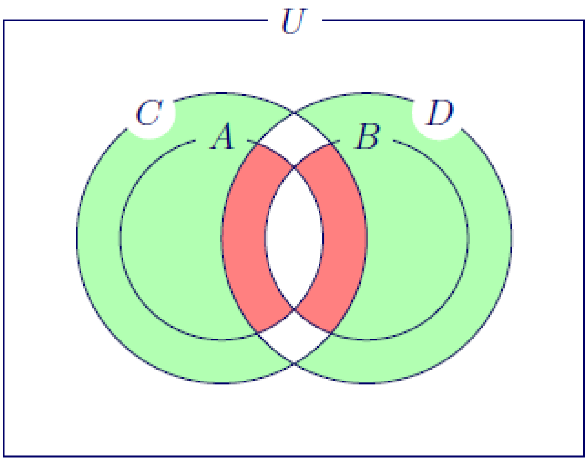
(a) Draw a Venn diagram showing four sets A, B, C, D in the most general configuration for which $A \subseteq C$ and $B \subseteq D$.

(b) By referring to your answer to (a), decide on the truth or falsity of the claim that, for all sets A, B, C, D,

$$[(A \subseteq C) \wedge (B \subseteq D)] \Rightarrow [(A \triangle B) \subseteq (C \triangle D)].$$

(c) [Challenge] Prove your answer to (b).

(a)



(b) Shown in red on the diagram are sections of $A \triangle B$ which are not part of $C \triangle D$ (shown green).

Thus even though $(A \subseteq C) \wedge (B \subseteq D)$ is true, $(A \triangle B) \subseteq (C \triangle D)$ is false.

$\boxed{\textit{So the claim is false.}}$

(c) An extreme counterexample occurs when $A \neq B$ but $C = D$. In this case $(A \triangle B) \neq \emptyset$ but $(C \triangle D) = \emptyset$ and so $(A \triangle B)$ cannot be a subset of $(C \triangle D)$.

To ensure that A is a subset of C and B is a subset of D we could take:

$$\boxed{a = \{a\},\ B = \{b\},\ a \neq b,\ C = D = \{a, b\}}$$

Week 3

Question 3 Let $A = \{x, y\}$ and let $B = \{4, 7, 9\}$.

1. Use set roster notation to list all possible functions from A to B. Give each of your functions a different name.
2. Use set roster notation to list all possible functions from B to A. Give each of your functions a different name.
3. Which, if any, of your functions from part 1 are injective?
4. Which, if any, of your functions from part 1 are surjective?
5. Which, if any, of your functions from part 2 are injective?
6. Which, if any, of your functions from part 2 are surjective?

1. The nine functions from A to B are:

$$
\begin{aligned}
f_1 &= \{(x,4),(y,4)\} & f_6 &= \{(x,7),(y,9)\}\\
f_2 &= \{(x,4),(y,7)\} & f_7 &= \{(x,9),(y,4)\}\\
f_3 &= \{(x,4),(y,9)\} & f_8 &= \{(x,9),(y,7)\}\\
f_4 &= \{(x,7),(y,4)\} & f_9 &= \{(x,9),(y,9)\}\\
f_5 &= \{(x,7),(y,7)\} & &
\end{aligned}
$$

2. The eight functions from B to A are:

$$
\begin{aligned}
g_1 &= \{(4,x),(7,x),(9,x)\} & g_5 &= \{(4,y),(7,x),(9,x)\}\\
g_2 &= \{(4,x),(7,x),(9,y)\} & g_6 &= \{(4,y),(7,x),(9,y)\}\\
g_3 &= \{(4,x),(7,y),(9,x)\} & g_7 &= \{(4,y),(7,y),(9,x)\}\\
g_4 &= \{(4,x),(7,y),(9,y)\} & g_8 &= \{(4,y),(7,y),(9,y)\}
\end{aligned}
$$

3. The functions $f_2, f_3, f_4, f_6, f_7, f_8$ are injective.

4. None of the functions are surjective.

5. None of the functions are injective

6. The functions $g_2, g_3, g_4, g_5, g_6, g_7$ are surjective.

Question 4 Prove or disprove each of the following statements:

1. There exists a bijection from $\mathbb{N}$ to $\mathbb{Z}$.
2. Every function from $\mathbb{Z}$ to $\mathbb{N}$ is surjective.
3. For any nonempty set A, any function from A to A is a bijection.
4. For any nonempty set A, there exists a bijection from A to A.

1. *The statement is true. For example, consider the function $f : \mathbb{N} \to \mathbb{Z}$ defined as follows*

n	1	2	3	4	5	...
$f(n)$	0	1	-1	2	-2	...

This function is surjective and injective, and hence a bijection.

2. *The statement is false. Consider the function $f : \mathbb{Z} \to \mathbb{N}$ defined by $f(z) = 12$. Note that Range $f = \{12\}$. Since Range $f \neq \mathbb{N}$, f is not surjective.*

3. *This is false. Consider $A = \{1, 2\}$ and the function $f : A \to A$ defined by $f(1) = f(2) = 1$. Then f is a function on A, but f is not surjective.*

4. *This is true. Let A be a nonempty set. The identity function $Id_A : A \to A$ defined by $Id_A(a) = a$ is a bijection on A.*

Question 5 Let $x = 11010_2,\ y = 10111_2,\ s = x + y,\ d = x - y,\ p = xy$.

(a) Calculate s, d and p directly in binary. Keep the answers in binary.

(b) Convert x, y to hexadecimal and recalculate s, d and p directly in hexadecimal.

(c) Finally convert x, y to decimal, recalculate s, d and p (in decimal) and convert the answers to hexadecimal and thence to binary.
Use the results to check your answers to (a) and (b).

(a)

$$\begin{array}{rcccccc} & & 1 & 1 & 0 & 1 & 0 \\ + & & 1 & 0 & 1 & 1 & 1 \\ \hline & 1 & 1 & 0 & 0 & 0 & 1 \\ \hline & {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 1} & & \end{array} \qquad \begin{array}{rccccc} & 1 & 1 & {}^{1}0 & {}^{1}1 & {}^{1}0 \\ - & 1 & 0_1 & 1_1 & 1_1 & 1 \\ \hline & 0 & 0 & 0 & 1 & 1 \\ \hline \end{array} \qquad \begin{array}{rccccccccc} & & & & & 1 & 1 & 0 & 1 & 0 \\ & & & & \times & 1 & 0 & 1 & 1 & 1 \\ \hline & & & & & 1 & 1 & 0 & 1 & 0 \\ & & & & 1 & 1 & 0 & 1 & 0 & 0 \\ & & & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline & {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 10} & {\scriptstyle 1} & {\scriptstyle 1} & & & & \end{array}$$

(b)

$$x = 11010_2 = \overbrace{0001}^{1}\overbrace{1010}^{A}{}_2 = 1A_{16}$$

$$y = 10111_2 = \overbrace{0001}^{1}\overbrace{0111}^{7}{}_2 = 17_{16}$$

$$\begin{array}{rcc} & 1 & A \\ + & 1 & 7 \\ \hline & 3 & 1 \\ \hline & {\scriptstyle 1} & \end{array} \qquad \begin{array}{rcc} & 1 & A \\ - & 1 & 7 \\ \hline & & 3 \\ \hline \end{array} \qquad \begin{array}{rccc} & & 1 & A \\ & \times & 1 & 7 \\ \hline & & B & 6 \\ & & {\scriptstyle (4)} & \\ & 1 & A & 0 \\ \hline & 2 & 5 & 6 \\ \hline & {\scriptstyle 1} & & \end{array}$$

(c)

$$x = 1A_{16} = (16 + 10)_{10} = 26_{10}$$

$$y = 17_{16} = (16 + 7)_{10} = 23_{10}$$

$$\begin{array}{rcc} & 2 & 6 \\ + & 2 & 3 \\ \hline & 4 & 9 \\ \hline \end{array} \qquad \begin{array}{rcc} & 2 & 6 \\ - & 2 & 3 \\ \hline & & 3 \\ \hline \end{array} \qquad \begin{array}{rccc} & & 2 & 6 \\ & \times & 2 & 3 \\ \hline & & 7 & 8 \\ & & {\scriptstyle (1)} & \\ & 5 & 2 & 0 \\ & {\scriptstyle (1)} & & \\ \hline & 5 & 9 & 8 \\ \hline \end{array}$$

$$49_{10} = (3 \times 16 + 1)_{10} = 31_{16} \checkmark = \underbrace{3}_{0011}\underbrace{1}_{0001} = 110001_2. \checkmark$$

$$3_{10} = 3_{16} \checkmark = 11_2. \checkmark$$

$$\begin{aligned} 598_{10} = (512 + 86)_{10} &= (2 \times 256 + 80 + 6)_{10} \\ &= (2 \times 16^2 + 5 \times 16 + 6)_{10} \\ &= 256_{16} \checkmark = \underbrace{2}_{0010}\underbrace{5}_{0101}\underbrace{6}_{0110} = 1001010110_2. \checkmark \end{aligned}$$

Question 6 This question is about the toggle-plus-one method as it relates to the storage of integers in computer words. It also shows how this method avoids the need for separate subtraction circuits.

As demonstrated in lectures, for a binary word W, toggle-plus-one means:

toggle:	replace every 1 by 0, every 0 by 1, then
add one:	treating W as a binary number, add 1.
	Ignore any carry beyond the length of the word.

For $l \in \mathbb{N}$ let $S_l = \{n \in \mathbb{Z} : -2^{l-1} \le n < 2^{l-1}\}$. Then S_l is the set of all integers that can be stored in words of length l bits, using the standard computer representation. The rules governing the storage of an integer n in a word W may be summarised as:

Rule 1:	n is negative if and only the left-most bit of W is 1
Rule 2:	$-n$ is stored as the word obtained from W by toggle-plus-one.
	This is true even when n is negative.
Rule 3:	If the left-most bit of W is 0 than $n = W_2$.
	i.e. n is retrieved by treating W as a binary number.

In computer arithmetic, subtraction uses negation and addition: $x - y = x + (-y)$.
In the addition, any carry beyond the length of the word is ignored.
As an example, take $x = 11010_2$, $y = 10111_2$ and use 8-bit computer arithmetic on:

(a) $x - y$ (b) $y - x$ (c) $-x - y$

Check your results by expressing x, y and your answers in decimal.

$$x = 11010_2 \longrightarrow \mathtt{00011010} \qquad x = 16 + 8 + 2 = 26_{10}.$$

$$-x \longrightarrow \mathtt{11100101}$$
$$+1$$
$$\boxed{\mathtt{11100110}}$$

$$y = 10111_2 \longrightarrow \mathtt{00010111} \qquad y = 16 + 4 + 2 + 1 = 23_{10}.$$

$$-y \longrightarrow \mathtt{11101000}$$
$$+1$$
$$\boxed{\mathtt{11101001}}$$

(a)

$$\begin{array}{rl} x: & \mathtt{00011010} \\ +(-y): & \mathtt{11101001} \\ (1) & \boxed{\mathtt{00000011}} \\ & \scriptstyle 1\,1\,1\,1 \end{array}$$

$$\longrightarrow 3_{10} = 26 - 23 \checkmark$$

(b)

$$\begin{array}{rl} y: & \mathtt{00010111} \\ +(-x): & \mathtt{11100110} \\ & \boxed{\mathtt{11111101}} \\ & \mathtt{00000010} \\ & \quad\quad\;\; +1 \\ \hline & \mathtt{00000011} \end{array}$$

$$\to -3_{10} = 23 - 26 \checkmark$$

(c)

$$\begin{array}{rl} -x: & \mathtt{11100110} \\ +(-y): & \mathtt{11101001} \\ (1) & \boxed{\mathtt{11001111}} \\ & \mathtt{00110000} \\ & \quad\quad\;\; +1 \\ \hline & \mathtt{00110001} \end{array}$$

$$-49_{10} = -26 - 23 \checkmark$$

Week 4

Question 4 Read the following algorithm:

Algorithm for converting a fraction x into binary with p binary places.
Inputs: fraction $x \in \mathbb{Q}$, $0 < x < 1$ and number of places $p \in \mathbb{N}$.
Initialise: $j \leftarrow 1$, $b_1, b_2, \ldots, b_p \leftarrow 0$
Loop: if $j = p + 1$ stop.
$x \leftarrow 2x$
If $x \geq 1$ $[\, b_j \leftarrow 1, \quad x \leftarrow x - 1 \,]$
$j \leftarrow j + 1$
Outputs: bits $b_1, b_2, \ldots, b_p$ such that $(0.b_1 b_2 \ldots b_p)_2$ is the best approximation to x using p binary places.

(a) Verify that the algorithm converts $\frac{1}{6}$ to 0.00101_2 when $p = 5$. Can you spot how to express $\frac{1}{6}$ exactly in repeating binary?

(b) Express 0.45_{10} as accurately as possible with 8 binary places.

(a)	x	j	b_{j-1}
initially	$\frac{1}{6}$	1	—
after loop 1	$\frac{1}{3}$	2	$0 = b_1$
2	$\frac{2}{3}$	3	$0 = b_2$
3	$\frac{1}{3}$	4	$1 = b_3$
4	$\frac{2}{3}$	5	$0 = b_4$
5	$\frac{1}{3}$	6	$1 = b_5$

stop

We see from the trace table above that the situation after the 4th loop is eactly the same as after the 2nd loop. So the group of the second and third digits repeats forever, and hence

$$\boxed{\frac{1}{6} = 0.0\overline{01}}.$$

(b)	x	j	b_{j-1}
initially	.45	1	—
after loop 1	.9	2	$0 = b_1$
2	.8	3	$1 = b_2$
3	.6	4	$1 = b_3$
4	.2	5	$1 = b_4$
5	.4	6	$0 = b_5$
6	.8	7	$0 = b_6$
7	.6	8	$1 = b_7$
8	.2	9	$1 = b_8$

stop

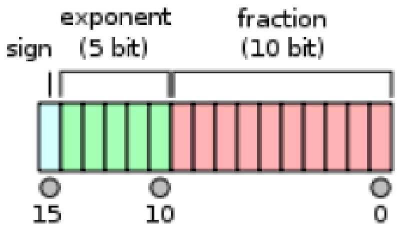
Hence to 8 binary places accuracy

$$\boxed{0.45_{10} \approx 0.01110011_2.}$$

Note that as with (a) it is easy to see the repeating pattern in the trace table and so deduce the exact repeating binary expression

$$0.45_{10} = 0.01\overline{1100}_2.$$

Question 5 The shortest IEEE standard for representing rational numbers is called *half-precision floating point*. It uses a 16-bit word partitioned as in the diagram at right. (This diagram is taken from the Wikipedia article on the subject, where more details can be found.)



As described in lectures, to store a rational number x it is first represented as $(-1)^s \times m \times 2^n$ with $1 \le m < 2$. The sign bit s is stored as the left-most bit (bit 15), the mantissa m (called "significand" in IEEE parlance) is stored in the right-most 10 bits (bits 9 to 0), and the exponent n is stored in the 5 bits in between (bits 14 to 10). However:

- *Only the fractional part of m is stored.* Because $1 \le m < 2$, the binary representation of m always has the form $1 \cdot \star\star\star\ldots$ where the stars stand for binary digits representing the fractional part of m. Hence there is no need to store the $1\cdot$ part.

- *the exponent n is stored with an "offset".* In order to allow for both positive and negative exponents, but to avoid another sign bit, the value stored is $n + 15$. In principle this means that the five exponent bits can store exponents in the range $-15 \le n \le 16$, but 00000 and 11111 are reserved for special purposes so in fact n is restricted to the range $-14 \le n \le 15$.

(a) A rational number x is stored in half-precision floating point as the word $3A6B_{16}$. (That's hex shorthand for the 16-bit binary word.) Write x in ordinary decimal notation.

(b) Find the word representing $x = -123.45_{10}$ in half-precision floating point. Use the closest approximation to x that it is possible to store in this format. Give your answer using hex shorthand, like the word you were given for (a).
Note: From a previous questions, $0.45_{10} = 0.01\overline{1100}_2$.

(a) $$3\mathrm{A}6\mathrm{B}_{16} = \underbrace{0011}_{3}\underbrace{1010}_{\mathrm{A}}\underbrace{0110}_{6}\underbrace{1011}_{\mathrm{B}} = \begin{array}{c|c|c} 0 & 01110 & 1001101011 \\ s & n+15 & \textit{frac.part of } m \\ & = 14 & \end{array}$$

$$\begin{aligned}
\longrightarrow\ & (-1)^0 \times 1.1001101011 \times 2^{14-15} \\
&= 1 \times 1.1001101011 \times 2^{-1} \\
&= (0.11001101011)_2 \qquad \textit{(moving the binary point 1 place to the left)} \\
&= \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{32} + \frac{1}{64} + \frac{1}{256} + \frac{1}{1024} + \frac{1}{2048}\right)_{10} \\
&= \frac{1024 + 512 + 64 + 32 + 8 + 2 + 1}{2048} = \frac{1643}{2048} \approx 0.802246_{10}.
\end{aligned}$$

$$
\begin{aligned}
(b) \qquad 123 &= -(64+32+16+8+2+1 \quad \textit{and} \quad 0.45 = 0.01\overline{1100}_2 \\
-123.45 &= -1111011.011100110011\ldots \\
&\approx -1.1110110111 \times 2^6 \qquad \textit{(shifting the binary point 6 places} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \textit{left and truncating)} \\
&\longrightarrow \begin{array}{c|c|l} \mathtt{1} & \mathtt{10101} & \mathtt{1110110111} \\ s & 21 & \textit{frac.part} \end{array} \qquad (6+15=21=\mathtt{10101}_2) \\
&\longrightarrow \underbrace{\mathtt{1101}}_{\text{D}}\underbrace{\mathtt{0111}}_{7}\underbrace{\mathtt{1011}}_{\text{B}}\underbrace{\mathtt{0111}}_{7} = \boxed{\text{D7B7}_{16}}
\end{aligned}
$$

Week 5

Question 4 I borrow \$100 000 at an interest rate of 6% per annum and agree to pay back \$1000 per month. Assuming interest compounding monthly, my debt, in dollars, after n months is given implicitly by

$$a_0 = 100\,000 = a \qquad a_n = (1.005)a_{n-1} - 1000 = ra_{n-1} - f$$

say, where $a = 100\,000$, $r = 1.005$ and $f = 1000$.

(a) Prove by mathematical induction that $\forall n \in \mathbb{N} \quad a_n = ar^n - f\sum_{k=0}^{n-1} r^k$.

(b) Given that for $r \neq 1$ and $n \in \mathbb{N}$, $\sum_{k=0}^{n-1} r^k = \frac{r^n - 1}{r - 1}$, how much will I owe in 10 years time?

(a) Let

$$P(n): \quad a_n = ar^n - f\sum_{k=0}^{n-1} r^k.$$

We will use mathematical induction to prove that $\forall n \in \mathbb{N}$ $P(n)$ ***Basis Step:***

$$\text{LHS of } P(1) = a_1 = ra_0 - f \quad \text{(using the inductive defn).}$$

$$\text{RHS of } (P(1) = ra^1 - f\sum_{k=0}^{1-1} r^k = ra - f\sum_{k=0}^{0} r^k = ra - fr^0 = ra - f.$$

Hence $P(1)$ holds.

Inductive step: *Let $n \in \mathbb{N}$. Suppose that $P(1), P(2), \ldots, P(n)$ all hold. We will show that $P(n+1)$ holds.*

$$\text{LHS of } P(n+1) = a_{n+1}$$

$$ra_n - f \qquad \text{(using the implicit definition)}$$

$$= r\left(ar^n - f\sum_{k=0}^{n-1} r^k\right) - f \qquad \text{(using } P(n))$$

$$= \qquad ar^{n+1} - f\left(\sum_{k=0}^{n-1} r^{k+1} + 1\right)$$

$$= ar^{n+1} - f\sum_{k=0}^{n} r^k = ar^{n+1} - f\sum_{k=0}^{(n+1)-1} r^k$$

Hence $P(n+1)$ holds

By the principle of mathematical induction, $P(n)$ holds for all $n \in \mathbb{N}$.

(b)

$$
\begin{aligned}
a_{120} &= 100\,000(1.005)^{120} - 1000\left(\frac{(1.005)^{120}-1}{1.005-1}\right)\\
&= 100\,000(1.005)^{120} - 200\,000((1.005)^{120}-1)\\
&= 100\,000(2-(1.005)^{120}) = \boxed{\textit{\$18\,060.33}}.
\end{aligned}
$$

Question 5 The letters of the word **TROUNCED** form the list $(X_i)_{1..8} = (\mathrm{T,R,O,U,N,C,E,D})$. This list is to be sorted into alphabetical order using Selection sort. The sorting is to be achieved by progressively modifying an index function π, rather than by shuffling members of the list itself. So initially

$$(X_i)_{1..8} = (X_{\pi(i)})_{1..8} \text{ where } \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{pmatrix}$$

and when sorting is complete π is sufficiently changed so that $(X_{\pi(i)})_{1..8}$ is in order.

(a) First apply the Least Element algorithm to $(X_i)_{1..8}$. Demonstrate the application by completing the trace table at right.

i	2	3	4	5	6	7	8	9
m	1	2	3	3	5	6	6	6
$x_{\pi(i)}$	R	O	U	N	C	E	D	-
$x_{\pi(m)}$	T	R	O	O	N	C	C	C

(b) Write out the modified index function π resulting from (a). $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 2 & 3 & 4 & 5 & 1 & 7 & 8 \end{pmatrix}$

(c) Now apply the Least Element algorithm to $(X_{\pi(i)})_{2..8}$ using this modified π, again demonstrating the application by a trace table.

(d) Write out the newly modified index function π resulting from (c).

(e) Without making trace tables, write out the state of index function π after each of the remaining applications of the Least element algorithm needed to complete the Selection sort of (T,R,O,U,N,C,E,D).

(f) What is the total number of comparisons used during this sort?

(g) By contrast, how many comparisons, in total, would be used to sort (T,R,O,U,N,C,E,D) using the Merge sort algorithm? To find out, carry out the Merge sort algorithm on (T,R,O,U,N,C,E,D) and carefully count the comparisons, remembering that when the Merge algorithm reaches a stage where one of its input lists is empty, it does not need any more comparisons to complete its task.

(c)

i	3	4	5	6	7	8	9
m	2	3	3	5	5	7	8
$x_{\pi(i)}$	O	U	N	T	E	D	–
$x_{\pi(m)}$	R	O	O	N	N	E	D

(d) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 3 & 4 & 5 & 1 & 7 & 2 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & \mathbf{7} & 4 & 5 & 1 & \mathbf{3} & 2 \end{pmatrix}$ [C D E U N T O R] $\rightarrow$ $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 7 & \mathbf{5} & \mathbf{4} & 1 & 3 & 2 \end{pmatrix}$ [C D E N U T O R] $\rightarrow$ $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 7 & 5 & \mathbf{3} & 1 & \mathbf{4} & 2 \end{pmatrix}$ [C D E N O T U R]

$\rightarrow$ $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 7 & 5 & 3 & \mathbf{2} & 4 & \mathbf{1} \end{pmatrix}$ [C D E N O R U T] $\rightarrow$ $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 7 & 5 & 3 & 2 & 1 & 4 \end{pmatrix}$ [C D E N O R T U]

(f) $8(7)/2 = \boxed{28}$.

(g)

```
(T) (R) (O) (U) (N) (C) (E) (D)
(R,T)  (O,U)  (C,N)  (D,E)   1+1+1+1=4
  (O,R,T,U)     (C,D,E,N)    3+3=6
      (C,D,E,N,O,R,T,U)      4
```

$\boxed{\textit{Total } 14}$.

Question 6 In lectures we saw how use the Merge sort algorithm to sort a sequence of length $n = 2^r$ into ascending order. In fact the algorithm can be applied to sequences of any length $n \in \mathbb{N}$. At each iteration the current sorted sub-sequences are merged in pairs as for the 2^r case but if there are an odd number of sub-sequences then the 'left over' one just joins, unchanged, the newly formed sub-sequences at the next iteration. This will mean that the merge algorithm will sometimes need to merge sequences of unequal lengths, but this causes no problems.

For example, if Merge sort is used to sort the letters of the word `PROVISIONAL` into alphabetical order then the subsequences at each stage will be:

after 0th iteration	`(P),(R),(O),(V),(I),(S),(I),(O),(N),(A),(L);`
after 1st iteration	`(P,R),(O,V),(I,S),(I,O),(A,N),(L);`
after 2nd iteration	`(O,P,R,V),(I,I,O,S),(A,L,N);`
after 3rd iteration	`(I,I,O,O,P,R,S,V),(A,L,N);`
after 4th iteration	`(A,I,I,L,N,O,O,P,R,S,V).`

(a) Apply the Merge sort algorithm to sort the letters of the word `APPROPRIATION` into alphabetical order, showing the results of each iteration as in the example above.

(b) How many comparison operations are used to merge sort `APPROPRIATION`? As in Q5, remember that when the merge algorithm reaches the stage where one of its input lists is empty, it does not need any more comparisons to complete its task. For example, for `PROVISIONAL` there are only 5 comparisons during the first iteration, 8 in the 2nd, 7 in the 3rd and 5 in the last.

(c) How many comparison operations would be used if `APPROPRIATION` were sorted using the Selection sort algorithm?

(a) & (b):

```
(A) (P) (P) (R) (O) (P) (R) (I) (A) (T) (I) (O) (N)
 (A,P)  (P,R)  (O,P)  (I,R)  (A,T)  (I,O)  (N)  6×1=6
  (A,P,P,R)*     (I,O,P,R)      (A,I,O,T)   (N)  3×3*=9
     (A,I,O,P,P,P,R,R)          (A,I,N,O,T)      7+3=10
        (A,A,I,I,N,O,O,P,P,P,R,R,T)              12
```

$$Total \;= 6 + 9 + 10 + 12 = \boxed{37}.$$

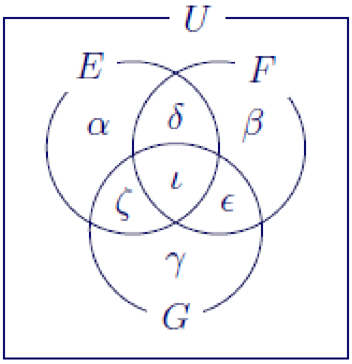
* *When comparing equal items, we 'take from the right'.*
This (arbitrary) choice is dictated by the particular way the psuedo code for the MergeSort algorithm was formulated in the lecture notes. See B2 Slide27.

(c) $13(12)/2 = \boxed{78}$.

Week 6

Question 4 (*Inclusion-exclusion principle and the product rule*)

(a) The inclusion-exclusion principle for two sets A, B is $|A \cup B| = |A| + |B| - |A \cap B|$. Use a Venn diagram to find a similar formula for three sets A, B, C.



In the diagram at left (taken from the relevant section of the lecture notes) the Greek letters represent the number of members of $E \cup F \cup G$ in each of its subsets created by intersections. So:

$$|E \cup F \cup G| = \alpha + \delta + \beta + \zeta + \iota + \epsilon + \gamma,$$
$$|E| = \alpha + \delta + \zeta + \iota, \qquad |E \cap F| = \delta + \iota,$$
$$|F| = \delta + \beta + \iota + \epsilon, \qquad |F \cap G| = \iota + \epsilon,$$
$$|G| = \zeta + \iota + \epsilon + \gamma, \qquad |G \cap E| = \zeta + \iota,$$
$$\text{and } |E \cap F \cap G| = \iota.$$

It follows that

$$\begin{aligned} |E|+|F|+|G| &= \alpha + 2\delta + \beta + 2\zeta + 3\iota + 2\epsilon + \gamma \\ &= |E \cup F \cup G| + \delta + \zeta + \epsilon + 2\iota \\ &= |E \cup F \cup G| + |E \cap F| + |F \cap G| + |G \cap E| - \iota. \end{aligned}$$

Hence

$$\boxed{|E \cup F \cup G| = |E|+|F|+|G| - |E \cap F| - |F \cap G| - |G \cap E| + |E \cap F \cap G|}.$$

(b) A PIN is a number with four decimal digits, *e.g.* 2357, 0944 etc. A 'double digit' in a PIN is any pair of consecutive equal digits, such as the 44 in 0944.

Use inclusion-exclusion to count how many PINs have at least one double digit.

[There is another, slightly quicker, way to count these PINs. Can you see it?]

Let L be the set of all PINs with a double digit at the left end.
Let R be the set of all PINs with a double digit at the right end.
Let M be the set of all PINs with a double digit in the middle.

We need $|L \cup M \cup R|$:

$$\begin{aligned} |L| &= |\{sstu : s,t,u \in \{0,\dots,9\}\}| = 10 \times 10 \times 10 = 10^3 \\ |M| &= |\{sttu : s,t,u \in \{0,\dots,9\}\}| = 10 \times 10 \times 10 = 10^3 \\ |R| &= |\{stuu : s,t,u \in \{0,\dots,9\}\}| = 10 \times 10 \times 10 = 10^3 \\ |L \cap M| &= |\{ssst : s,t \in \{0,\dots,9\}\}| = 10 \times 10 = 10^2 \\ |M \cap R| &= |\{sttt : s,t \in \{0,\dots,9\}\}| = 10 \times 10 = 10^2 \\ |L \cap R| &= |\{sstt : s,t \in \{0,\dots,9\}\}| = 10 \times 10 = 10^2 \\ |L \cap M \cap R| &= |\{ssss : s \in \{0,\dots,9\}\}| = 10. \end{aligned}$$

So $|L \cup M \cup R| = 3(10^3) - 3(10^2) + 10 = \boxed{2710}$.

An alternative to the above is to use complementary counting. Without restictions there are 10^4 PINs. With the restiction of no *double digits there are* $10 \times 9 \times 9 \times 9 = 7290$ *PINs. So the number* with *double digits is* $10000 - 7290 = 2710$.

Question 5 (*Combinations and 'stars and bars'*) A TAW is a three letter 'word' whose letters are in alphabetical order. Letters are drawn from the standard lower case English 26-letter alphabet, and are allowed to repeat. A 'word' does not have to appear in any dictionary but must contain at least one vowel and at least one consonant. Letter y can count as a vowel or a consonant. Examples of TAWs are *abc*, *ccy*, *aoy* and *yyy*.

How many different TAWs are there?

Hint 1: Once three letters are chosen (possibly involving repeats) there is only one way to put them in alphabetical order.
Hint 2: As a first step ignore the requirement about vowels and consonants.

From the hints, we first need to count the multisets of size 3 *with members from the alphabet. (Recall that a "multiset" is an un-ordered list with repeats allowed.) We use 'stars and bars' for this.*

Take 26 buckets, one for each letter of the alphabet, and drop three stars into this collection. We need 25 bars to separate the buckets.

For example, $|||\star||||\star\star\underbrace{||\cdots||}_{17 \text{ bars}}$ *represents* dii.

So the number of these multisets is $\binom{25+3}{3} = \binom{28}{3}$.

Now, using complementary counting, we must subtract the number of all-vowel 'words' and the number of all-consonant words.
We again use 'stars and bars', but with less buckets.

There are 5^* *vowels, so* $\binom{4+3}{3} = \binom{7}{3}$ *all-vowel words.*

There are 20^* *consonants, so* $\binom{19+3}{3} = \binom{22}{3}$ *all-consonant words.*

Thus the number of TAWs is

$$\begin{aligned}\binom{28}{3} - \binom{7}{3} - \binom{22}{3} &= \frac{28 \cdot 27 \cdot 26}{3.2.1} - \frac{7 \cdot 6 \cdot 5}{3.2.1} - \frac{22 \cdot 21 \cdot 20}{3.2.1} \\ &= 3276 - 35 - 1540 = \boxed{1701}.\end{aligned}$$

* *Since letter* y *is ambivalent, we do not count it amongst the vowels or the consanents when counting all-vowel and all-consonant words. For example we do not want to eliminate* eye *nor* cry.

Question 6 (*The pigeon hole principle*)

(a) Each year the ANU enrols new students from all eight States and Territories in Australia. How many new Australian students must the ANU enrol next year in order to ensure that there are at least 500 students from the same State or Territory?

The students are the pigeons; the States and Territories are the pigeon holes.

We need the smallest n for which $\left\lceil \dfrac{n}{8} \right\rceil = 500.$

So take $n = 8 \times 499 + 1 = \boxed{3993}$.

(b) Seven numbers are picked from the set $\{1, 2, \ldots, 12\}$ of the first twelve natural numbers. Prove that amongst the seven numbers picked it is guaranteed that two of them, say a and b satisfy $a - b = 3$.

Remark: This is easy once you've figured out what the pigeon holes should be, but figuring that out is perhaps not so easy!

Make six pigeon holes, each labelled with ***two*** *numbers, like this:*

1,4	*2,5*	*3,6*	*7,10*	*8,11*	*9,12*

Then every number in $\{1, 2, \ldots, 12\}$ is part of exactly one label.

The given seven numbers are the pigeons and each roosts in the hole with its number on it.

With seven pigeons and only six holes, at least one hole contains two pigeons. The corresponding numbers differ by 3.

[This question is from the preparatory question set for section C1.]

Week 7

Question 1 In rôle-playing games, such as the original, *Dungeons and Dragons*, gamers use a variety of dice shapes, not just the standard cube. For example a 'd12' has the shape of a regular dodecahedron. A (fair) d12 has twelve faces numbered 1,2,. . . ,12, and each face is equally likely to become the top face (the face that is read) when the die is 'thrown'.

Suppose gamer Alice does not have a d12 but does have a d6 (*i.e.* a regular cubic die with six faces labelled 1,2,. . . ,6.) To simulate throwing a d12, Alice throws her d6 twice, resulting in a pair of values $(a, b) \in \{1, \dots, 6\}^2$. She then combines the two values in some way to come up with a value in $\{1, \dots, 12\}$.

(a) Write out the sample space S for this 'experiment'.

$$\left\{\begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3) & ,(4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array}\right\}$$

(b) One way to combine a and b would be to add them. But this would not accurately simulate a d12 throw because it would be impossible to score 1.
To avoid this problem Alice decides to use the formula: $v = \left\lceil \frac{ab}{3} \right\rceil$.
Write out the event $E \subseteq S$ coresponding to $v = 1$.

We need $ab \leq 3$ *so*

$$E = \{(1,1),\ ((1,2),\ (1,3),\ (2,1),\ (3,1)\}.$$

(c) What is the probability of event E above and in what way does it demonstrate that Alice's method also does not accurately simulate a d12 throw?

$$\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{5}{36}.$$

On a d12 every outcome has equal probability $\frac{1}{12}$. *But* $\frac{5}{36} \neq \frac{1}{12}$.

(d) How *could* Alice combine a and b to accurately simulate a d12 throw?
(Gamers actually do this, so you can Google an answer if you're desperate!)

One possibility is:

$$v = \begin{cases} a & \text{if } b \text{ is even} \\ a+6 & \text{if } b \text{ is odd} \end{cases}$$

For example, $\mathbb{P}(v=5) = \frac{1}{36}|\{(5,2),(5,4),(5,6)\}| = \frac{3}{36} = \frac{1}{12}$.

Question 4 The income and education level of each person on the electoral roll for Queanberra is recorded as a pair $(x, y) \in \{1, 2, 3\}^2$, where 1 stands for low, 2 for average, and 3 for high, *e.g.* $(2, 3)$ represents a highly educated person with average income.

Let S denote the set of all people on the Queanberra electoral roll, and define random variables $X, Y : S \to \{1, 2, 3\}$ by $X(s), Y(s)$ are the income and educational levels of person s. Let $p_{i,j} = \mathbb{P}(\{(X(s) = i) \wedge (Y(s) = j)\})$ for $1 \leq i, j \leq 3$. Assume that

$$(p_{i,j})_{1 \leq i,j \leq 3} = \begin{pmatrix} 0.05 & 0.10 & 0.05 \\ 0.10 & 0.20 & 0.10 \\ 0.15 & 0.20 & 0.05 \end{pmatrix}.$$

(a) Explain what $p_{1,2}$ represents.

$\boxed{\textit{Probability that } s \textit{ has low income and average education level.}}$

(b) Find the probability of the event AI: "the person has an average income"?

$p_{2,1} + p_{2,2} + p_{2,3} = 0.10 + 0.20 + 0.10 = \boxed{0.40}.$

(c) Find the probability of the event AE: "the person has an average education level"?

$p_{1,2} + p_{2,2} + p_{3,2} = 0.10 + 0.20 + 0.20 = \boxed{0.50}.$

(d) Describe the event $AI \cup AE$ and find its probability.

$\boxed{s \textit{ has average income or average education level}}$

$$\begin{aligned} \mathbb{P}(AI \cup AE) &= \mathbb{P}(AI) + \mathbb{P}(AE) - \mathbb{P}(AI \cap AE) \\ &= 0.40 + 0.50 - p_{2,2} \;=\; 0.90 - 0.20 \;=\; \boxed{0.70}. \end{aligned}$$

(e) Are the events AI and AE independent?

$\mathbb{P}(AI) \times \mathbb{P}(AE) = 0.40 \times 0.50 = 0.20 = p_{2,2} = \mathbb{P}(AI \cap AE).$

So $\boxed{\textit{Yes}}$ *they are independent.*

(f) Are the random variables X and Y independent?

$\boxed{\textit{No}}$. *E.g. let* LI = *'Low Income' and let* LE = *'Low Ed. level'. Then*

$$\begin{aligned} \mathbb{P}(X(1)) \times \mathbb{P}(Y(1)) &= \mathbb{P}(LI) \times \mathbb{P}(LE) \\ &= (0.05{+}0.10{+}0.05{+}) \times (0.05{+}0.10{+}0.15) \\ &= 0.20 \times 0.30 = 0.06 \neq 0.05 = p_{1,1} = \mathbb{P}(X(1) \cap Y(1)). \end{aligned}$$

Week 8

Question 1 Recall the following from our lecture notes.

Lemma:
For any probability experiment with sample space S, and for any events $A, B \subseteq S$, if $\mathbb{P}(A) \neq 0$ then

$$\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A).$$

Theorem:[Bayes' Theorem]
For any probability experiment with sample space S, for any $n \in \mathbb{N}$, for any partition $\{B_1, B_2, \ldots, B_n\}$ of S and for any event $A \subseteq S$, if $\mathbb{P}(A) \neq 0$ and for all $i \in \{1, 2, \ldots, n\}$ we have $\mathbb{P}(B_i) \neq 0$, then for all $k \in \{1, 2, \ldots, n\}$ we have

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

(i) Write out the structural part of a proof of the lemma that proceeds directly.

(ii) Complete the proof of the lemma you started in part (i).
HINT: Use the definition of conditional probability.

(iii) Write out the structural part of a proof of Bayes' Theorem that proceeds directly.

(iv) In our lecture proof of Bayes' Theorem, we wrote a sequence of 8 equalities to establish that, under the hypotheses made, the conclusion of Bayes' Theorem holds. Each equality was accompanied by a justification. Below is a table of the 8 justifications in the order in which they appeared in our lecture proof of Bayes' Theorem. Use these to finish the proof of Bayes' Theorem you started in part(ii).

#	Justification for algebraic manipulation
1	(By defn of $\mathbb{P}(B_k\|A)$)
2	(Applying the lemma, which is OK because $\mathbb{P}(B_k) \neq 0$)
3	(Because $A \cap S = A$)
4	(Because $\{B_1, \ldots, B_n\}$ is a partition of S, we have $S = B_1 \cup B_2 \cup \cdots \cup B_n$)
5	($\cap$ distributes over $\cup$)
6	(Applying the sum rule, which is OK because $B_1, \ldots, B_n$ are mutually disjoint)
7	(Applying the lemma n times, which is OK because $\mathbb{P}(B_i) \neq 0$ for $i \in \{1, 2, \ldots, n\}$)
8	(Using Σ notation)

Solution to Q1

(i) **Proof:** *Consider a probability experiment with sample space* S*. Let* $A, B \subseteq S$*. Suppose that* $\mathbb{P}(A) \neq 0$*.*

$\vdots$

Hence $\mathbb{P}(B \mid A)\mathbb{P}(A) = \mathbb{P}(A \cap B)$*.* $\square$

(ii) **Proof:** *Consider a probability experiment with sample space* S*. Let* $A, B \subseteq S$*. Suppose that* $\mathbb{P}(A) \neq 0$*. Since* $\mathbb{P}(A) \neq 0$*, the conditional probability* $\mathbb{P}(B|A)$ *is defined. The definition gives*

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

Multiplying both sides by $\mathbb{P}(A)$ *gives* $\mathbb{P}(B|A)\mathbb{P}(A) = \mathbb{P}(A \cap B)$*.* $\square$

(iii) **Proof:** *Consider a probability experiment with sample space* S*. Let* $n \in \mathbb{N}$*, let* $\{B_1, B_2, \ldots, B_n\}$ *be a partition of* S *and let* $A \subseteq S$*. Suppose that* $\mathbb{P}(A) \neq 0$ *and for all* $i \in \{1, 2, \ldots, n\}$ *we have* $\mathbb{P}(B_i) \neq 0$*. Let* $k \in \{1, 2, \ldots, n\}$*.*

$\vdots$

Hence

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)} \quad \square$$

(iv) **Proof:** *Consider a probability experiment with sample space S. Let $n \in \mathbb{N}$, let $\{B_1, B_2, \ldots, B_n\}$ be a partition of S and let $A \subseteq S$. Suppose that $\mathbb{P}(A) \neq 0$ and for all $i \in \{1, 2, \ldots, n\}$ we have $\mathbb{P}(B_i) \neq 0$. Let $k \in \{1, 2, \ldots, n\}$. Now*

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)} \tag{\#1}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)} \tag{\#2}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A \cap S)} \tag{\#3}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A \cap (B_1 \cup B_2 \cup \cdots \cup B_n))} \tag{\#4}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}((A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n))} \tag{\#5}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_1) + \cdots + \mathbb{P}(A \cap B_n)} \tag{\#6}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2) \cdots + \mathbb{P}(A|B_n)\mathbb{P}(B_n)} \tag{\#7}$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)} \tag{\#8} \qquad \square$$

Question 2 [This problem is Problem 14 in Excercise set 9.9 of Epp. (2019). Discrete Mathematics with Applications, metric Edition. Cengage. (our optional text)] A drug-screening test is used in a large population of people of whom 4% actually use drugs. Suppose that the false positive rate is 3% and the false negative rate is 2%. Thus a person who uses drugs tests positive for them 98% of the time, and a person who does not use drugs tests negative for them 97% of the time

(i) What is the probability that a randomly chosen person who tests positive for drugs actually uses drugs?

(ii) What is the probability that a randomly chosen person who tests negative for drugs does not use drugs?

(i) We consider an experiment in which a member of the population is selected at random and then tested for drugs. The outcome of the experiment is the person chosen, which we represent by their name, and the sample space S is the set of all names of people in the population. Since the person is selected at random, each outcome is equally likely, and for any event $E \subset S$, we may compute

$$\mathbb{P}(E) = \frac{|E|}{|S|}.$$

Let P be the event that the person tests positive for drugs, let N be the event that the person tests negative for drugs, let U be the event that the person uses drugs and let D be the event that the person does not use drugs. We note that $\{D, U\}$ is a partition of S. The data we are given may now be expressed in terms of the events we have introduced and conditional probabilities. We have:

$$\begin{aligned}
\mathbb{P}(P|U) &= 0.98 && \textit{(this is a true positive)}\\
\mathbb{P}(P|D) &= 0.03 && \textit{(this is a false positive)}\\
\mathbb{P}(N|U) &= 0.02 && \textit{(this is a false negative)}\\
\mathbb{P}(N|D) &= 0.97 && \textit{(this is a true positive)}
\end{aligned}$$

Since 4% of the population uses drugs, we have

$$\mathbb{P}(U) = 0.04 \textit{ and } \mathbb{P}(D) = 0.96$$

We compute

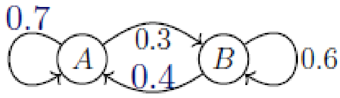
$$\begin{aligned}
&\textit{The probability that a randomly chosen person who tests positive} \\
&\qquad\qquad \textit{for drugs actually uses drugs} \\
=&\mathbb{P}(U|P) \quad \textit{(Converting to notation)} \\
=&\frac{\mathbb{P}(P|U)P(U)}{\mathbb{P}(P|U)P(U)+\mathbb{P}(P|D)P(D)} \quad \textit{(By Bayes' Theorem)} \\
=&\frac{0.98\times 0.04}{0.98\times 0.04+0.03\times 0.96} \quad \textit{(Using the data above)} \\
=&0.576 \\
=&57.6\%
\end{aligned}$$

(ii) Using the setup and notation from our solution to (i), we compute

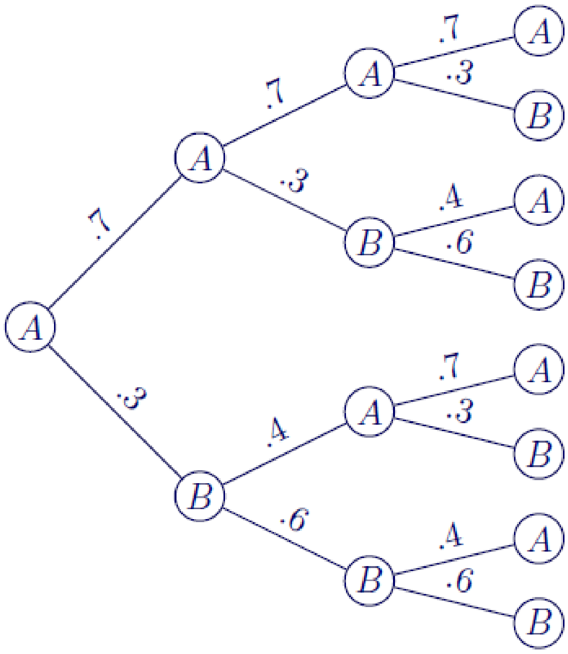
$$\begin{aligned}
&\textit{The probability that a randomly chosen person who tests negative} \\
&\qquad\qquad \textit{for drugs does not use drugs} \\
=&\mathbb{P}(D|N) \quad \textit{(Converting to notation)} \\
=&\frac{\mathbb{P}(N|D)P(D)}{\mathbb{P}(N|U)P(U)+\mathbb{P}(N|D)P(D)} \quad \textit{(By Bayes' Theorem)} \\
=&\frac{0.97\times 0.96}{0.97\times 0.96+0.02\times 0.04} \quad \textit{(Using the data above)} \\
=&0.999 \\
=&99.9\%
\end{aligned}$$

Question 3 A Markov process has two states A and B with transition graph below.

(a) Write in the two missing probabilities.



(b) Suppose the system is in state A initially. Use a tree diagram to find the probability that the system will be in state B after three steps.



$.7\times.7\times.7=0.343$

$.7\times.7\times.3=\mathbf{0.147}$

$.7\times.3\times.4=0.084$

$.7\times.3\times.6=\mathbf{0.126}$

$.3\times.4\times.7=0.084$

$.3\times.4\times.3=\mathbf{0.036}$

$.3\times.6\times.4=0.072$

$.3\times.6\times.6=\mathbf{0.108}$

Probability system in state B after 3 steps $= \boxed{0.417}$

(c) The transition matrix for this process is $T = \boxed{\begin{bmatrix} .7 & .3 \\ .4 & .6 \end{bmatrix}}$.

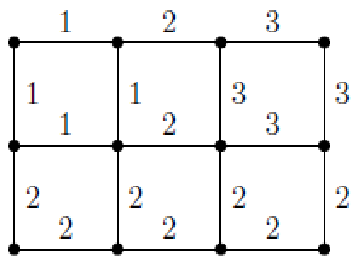
(d) Use T to recalculate the probability found in (b).

$$\begin{aligned}
\mathbf{x}_3 = (T')^3\mathbf{x}_0 &= \begin{bmatrix} .7 & .4 \\ .3 & .6 \end{bmatrix}\begin{bmatrix} .7 & .4 \\ .3 & .6 \end{bmatrix}\begin{bmatrix} .7 & .4 \\ .3 & .6 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} .7 & .4 \\ .3 & .6 \end{bmatrix}\begin{bmatrix} .7 & .4 \\ .3 & .6 \end{bmatrix}\begin{bmatrix} .7 \\ .3 \end{bmatrix} \\
&= \begin{bmatrix} .7 & .4 \\ .3 & .6 \end{bmatrix}\begin{bmatrix} .61 \\ .39 \end{bmatrix} \\
&= \begin{bmatrix} .583 \\ .417 \end{bmatrix}
\end{aligned}$$

So we again get: Probability system in state B after 3 steps $= \boxed{0.417}$.

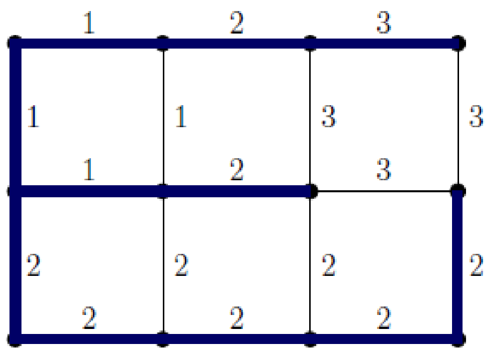
Week 9

Question 4 A weighted graph G is shown at right.



(a) Find a minimal spanning tree for G and calculate its total weight.

Trace out your answer below:



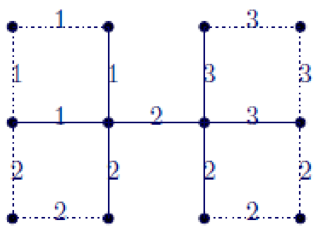
This is just one of many possible answers.

Total weight
$= 3\times 1 \; + \; 7\times 2 \; + \; 1\times 3$
$= 3 + 14 + 3 = \boxed{20}$

(b) Prove or disprove that G has a minimal spanning tree containing two vertices of degree 4.

Only the two central vertices of G have degree 4 and there are no vertices of higher degree.

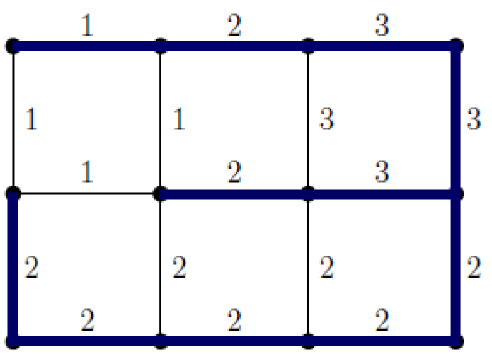
So a spanning tree with two vertices of degree 4 must include these two central vertices and all the edges incident on them. This contributes $1+1+2+2+2+3+3 = 14$ to the total weight.



Then to include the corner vertices, no matter which joining edges we choose, we must add another $1+3+2+2 = 8$ to the total weight.

Thus the total weight is $14+8 = 22 > 20$, and so the tree is not minimal.

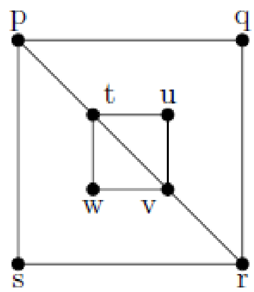
(c) Find a maximal spanning tree for G and calculate its total weight.
Trace out your answer below:



Using the minimimal spanning tree algorithm (Kruskal's), but picking maximal available edges instead of minimal ones, gives a maximal spanning tree. One of many is shown.

Total weight $= 3\times 3 \; + \; 7\times 2 \; + \; 1\times 1$
$= 9 + 14 + 1 = \boxed{24}$.

Question 6 For the graph J at right:



(a) Prove that J has no Euler circuit.

(b) Prove that J has no Hamilton circuit.

(c) Suppose that Fleury's algorithm is used to find an Euler path from p to r. What feature of the algorithm prevents the path starting ptvr?

1. *We have a theorem that says: A connected graph has an Euler circuit if and only if every vertex has even degree. Since p has degree 3, J has no Euler circuit.*

2. **We shall use a proof by contradiction. Suppose that J has a Hamilton circuit.** *Then J has an Euler circuit that starts and ends at w. Since J has 8 vertices, the Hamilton circuit has 9 vertices; the first and the ninth vertices are w. Since the only neighbours of w are t and v, the second vertex in the circuit must be t or v and the second last vertex in the circuit must be t or v. Since t and v are each visited once by the circuit, either t is the second vertex visited and v is the eighth vertex visited, or v is the second vertex visited and t is the eighth vertex visited. In the latter case, we may reverse the circuit and we still have a Hamilton circuit. So, without loss of generality, we may assume that t is the second vertex visited and v is the eighth vertex visited; that is, the circuit reads $wt \ldots vw$.*

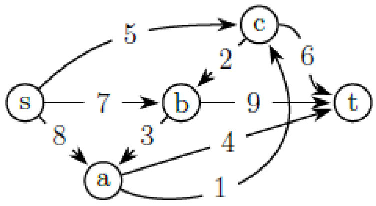
   The only neighbours of u are t and v. It follows that the circuit must visit u immediately after visiting one of t or v. Since, in this case, the circuit returns to w immediately after visiting v, we have that the circuit must visit u immediately after visiting t. That is, the circuit reads $wtu \ldots vw$. Since t is the second vertex and u is the third vertex in the circuit, and the only neighbours of u are t and v, it must be that v is the fourth vertex in the circuit. So the circuit reads $wtuv \ldots vw$. But then v is the fourth vertex and the eighth vertex visited in the circuit. This contradicts the fact that the circuit is a Hamilton circuit.

   Since our supposition that J has a Hamilton circuit has led to a contradiction, it cannot hold. $\square$

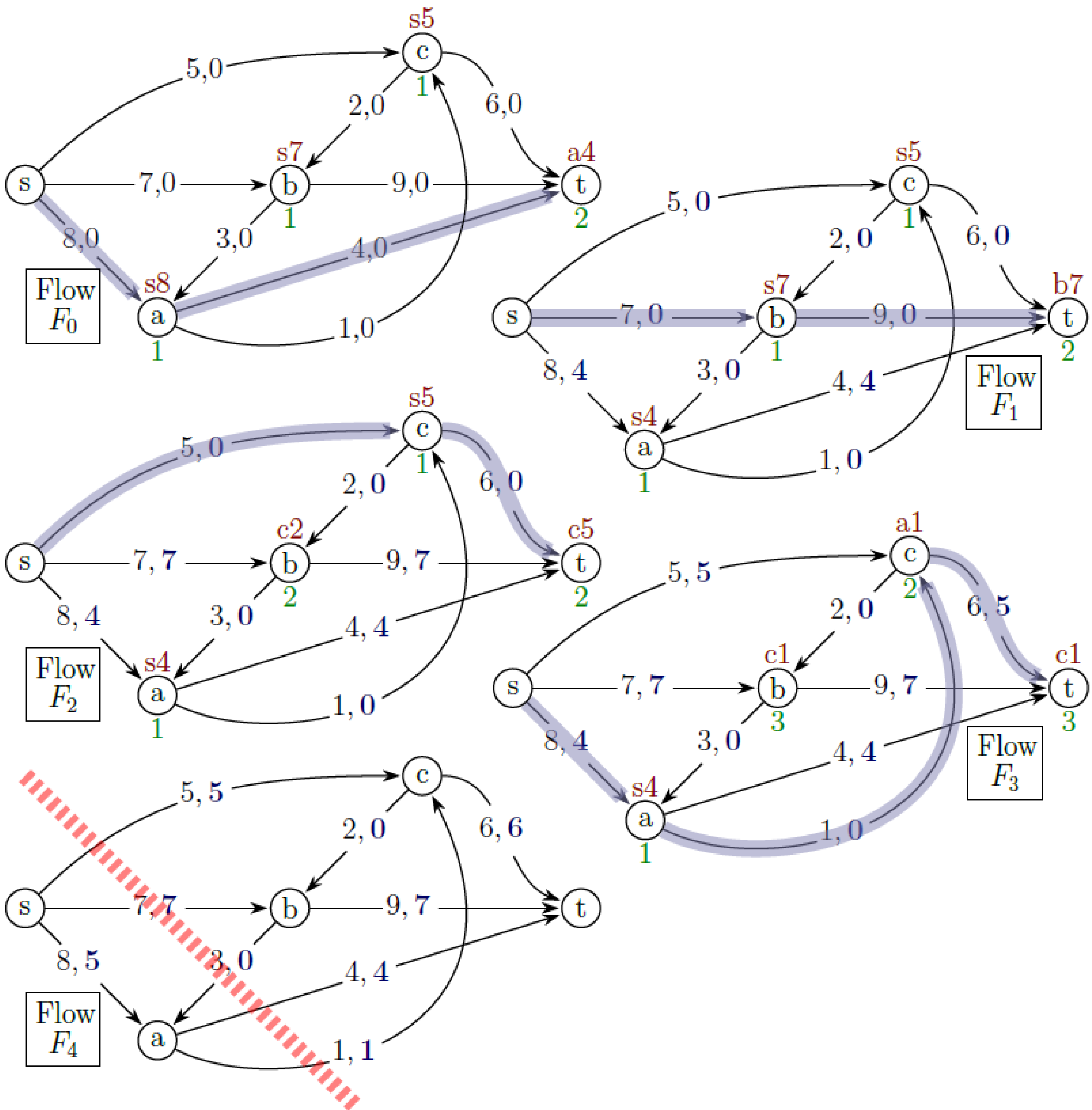
3. *If Fleury's algorithm began by constructing a path ptv, then the edge $\{v, r\}$ is a bridge in the remaining graph and would not be chosen as the next edge.*

Week 10

Question 1 The diagram at right shows the capacities and directions of all links in a network with source s, target t and intermediate nodes a, b and c. Use the labelling algorithm to find the maximum flow through the network and how it can be achieved. Prove that your flow is maximum by finding a cut of equal value.



Here are some blank diagrams to fill in with **levels, labels** and **flows**. Use an outliner pen to mark the **incremental flow** dictated by the target's label. The first diagram is filled in as an example. You will need all diagrams. Draw a min cut on the last diagram.



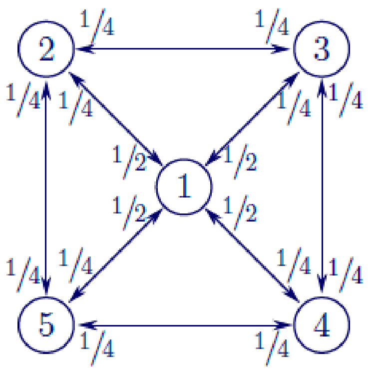
Draw a minimum cut on this diagram.

Max Flow Value = Min Cut Capacity = $\boxed{17}$

Question 4 Let G be the digraph with adjacency matrix A shown at right. A random walker on G has all transition probabilities equal to $^1/_4$ except that all transition to vertex 1 have probability $^1/_2$.

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

(a) Draw the graph and mark on it the transition probabilities.



(b) Compile the transition matrix T and verify that it is stochastic.

$$T = \begin{bmatrix} 0 & ^1/_4 & ^1/_4 & ^1/_4 & ^1/_4 \\ ^1/_2 & 0 & ^1/_4 & 0 & ^1/_4 \\ ^1/_2 & ^1/_4 & 0 & ^1/_4 & 0 \\ ^1/_2 & 0 & ^1/_4 & 0 & ^1/_4 \\ ^1/_2 & ^1/_4 & 0 & ^1/_4 & 0 \end{bmatrix}.$$

All entries are non-negative and all rows sum to 1.
So the matrix is stochastic.

(c) On average over the long term, what proportion of time will the walker spend at vertex 1? Hint: There are really only two unknowns.

Let the required probability be p.

By symmetry the long term probabilities for all the other vertices will be equal, say q. *Then* $S = \begin{bmatrix} p & q & q & q & q \end{bmatrix}'$.

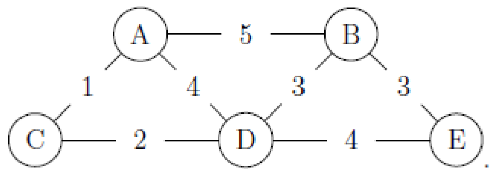
We need to solve $T'S = S$ *subject to* S *being a probability vector:*

First row of $(T'-I)S=0$	:	$-p+\frac{1}{2}q+\frac{1}{2}q+\frac{1}{2}q+\frac{1}{2}q = 0$[1]
Probability vector	:	$p+q+q+q+q = 1$
	$\Longrightarrow$	$6q = 1 \implies q = {}^1/_6 \implies p = {}^1/_3$

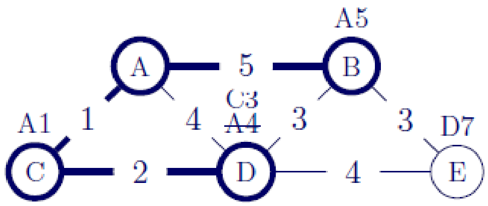
So the walker will spend $\boxed{\textit{one third}}$ *of the time at vertex* 1.

[1]This is equivalent to Prob(in)=Prob(out) at vertex 1.

Question 7 Let G be the weighted graph



(a) Show that Dijkstra's algorithm correctly finds that the shortest path from A to B is the direct path using just the edge AB. You can do all the work on a single diagram, but, to show that you have used the algorithm correctly, if an annotation needs updating do not erase it — just put a line through it and write the new annotation above that.



(b) In what order are the vertices added to the tree?

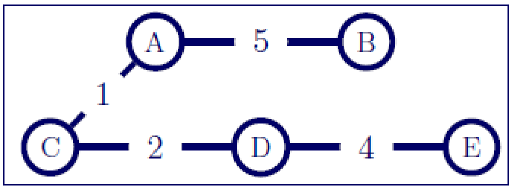
$\boxed{\text{(A), C, D, B}}$.

(c) Notice that the algorithm does not, in this instance, generate a spanning tree. Which vertex or vertices are missing?

$\boxed{\text{E}}$.

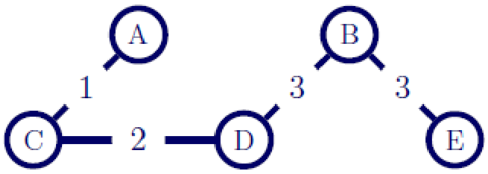
(d) Extend the use of the algorithm until the shortest distance from A to each other vertex is established, and a spanning tree is thereby generated. Draw this tree.

Vertex E is marked with D, so add edge DE to give the spanning tree:



(e) Is the spanning tree you have generated using Dijkstra's algorithm a minimal spanning tree? Justify your answer.

$\boxed{\textit{No}}$. *Kruskal's algorithm gives*



which has total weight $1+2+3+3=9$. *This total is less than that for the Dijkstra tree, which has total weight* $1+2+4+5=12$.