# COMS20010 — Exam

This exam was actually given as a Blackboard test in a slightly different format to the one used this year (hence "Section 1" and "Section 2"). This means this isn't quite the final version, and since it hadn't gone through the final round of checking at that point there might still be one or two errors lurking in the model answers, but it should be a good guide as to what to expect.

Section 1

1. (5 marks) (Short question.) Mark each of the following statements true or false:
   (a) (1 mark) $n \in O(4^n)$.
   (b) (1 mark) $\log n \in O(n^{1/3})$.
   (c) (1 mark) $n^{1.01} + 10^6 n \notin \omega(n^{1.01})$.
   (d) (1 mark) $mn^2 + m^2 n \in O(m^2 n^2)$ (where neither m nor n is a constant).
   (e) (1 mark) In any graph $G = (V, E)$, $|E| \in o(|V|^2)$.

2. (5 marks) (Short question.) Consider the following graphs G_1, G_2, and G_3, and the indicated vertex v.
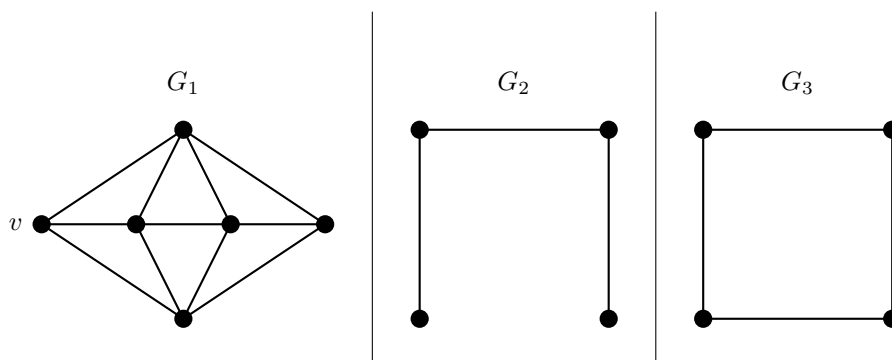

   Mark each of the following statements true or false:
   (a) (1 mark) G_1 is connected.
   (b) (1 mark) G_3 contains a Hamilton cycle.
   (c) (1 mark) G_1 contains an Euler walk from v back to itself.
   (d) (1 mark) G_1 contains a copy of G_2 as a (not necessarily induced) subgraph.
   (e) (1 mark) G_1 contains a copy of G_3 as an induced subgraph.

3. (5 marks) (Short question.) Consider the following flow network and flow, with source s and sink t.

a 4/6 b

5/8   4/4   5/5   3/4   7/7 t

s 2/4   4/8

2/5

c d
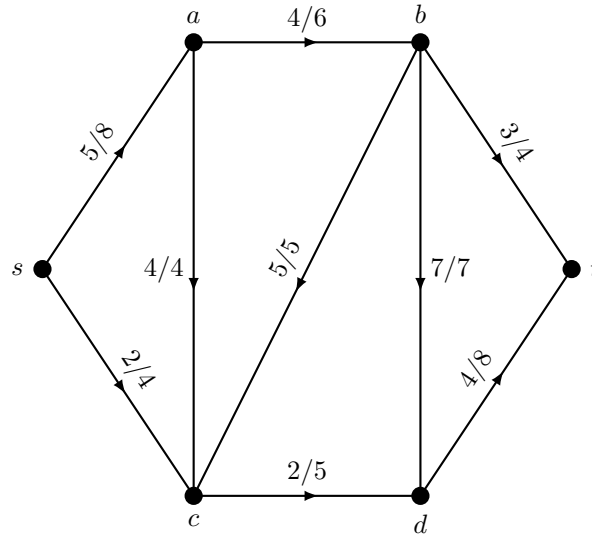
List all augmenting paths of the network under the given flow.

4. (5 marks) (Short question.) In the distant future, the renowned Sol Ballet Corps is putting on a production to emphasise a spirit of love and togetherness between human colonies throughout the solar system. They have 4 dancers from Earth, 5 from Mars, 6 from the moon, and 7 from Ceres. They wish to arrange a procession in which every pair of dancers **from different worlds** crosses the stage exactly once. For example, one dancer from Earth will cross alongside one dancer from Mars, but not a pair of dancers both from Earth. How many pairs of dancers will need to cross the stage in total? (**Hint:** Try modelling this as a graph and applying the handshaking lemma.)

    A. 179.
    B. 358.
    C. 175.
    D. 350.
    E. None of the above.

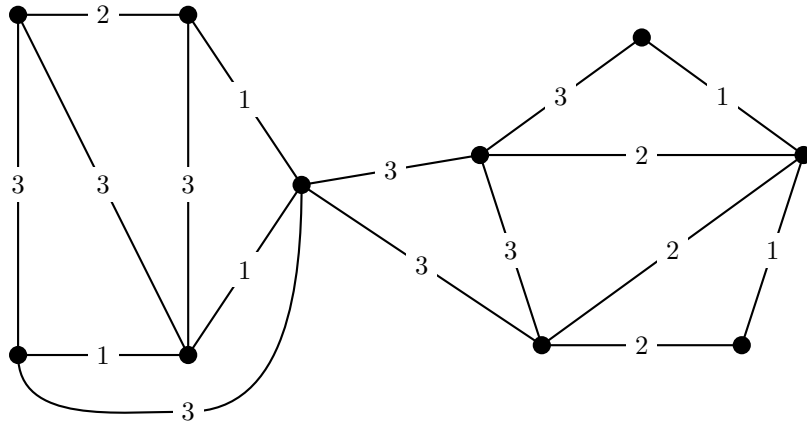5. (5 marks) (Short question.) Consider the weighted graph below.

20

7   2   8

4   10

u 9   5 v

8   6

3   1

9
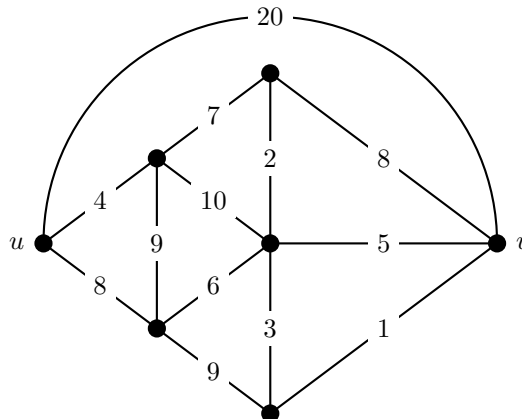
What is the distance from u to v?

    A. 16.
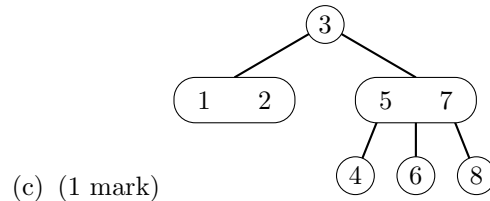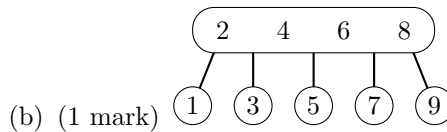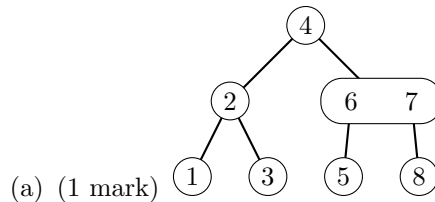
B. 17.

C. 18.

D. 19.

E. None of the above.

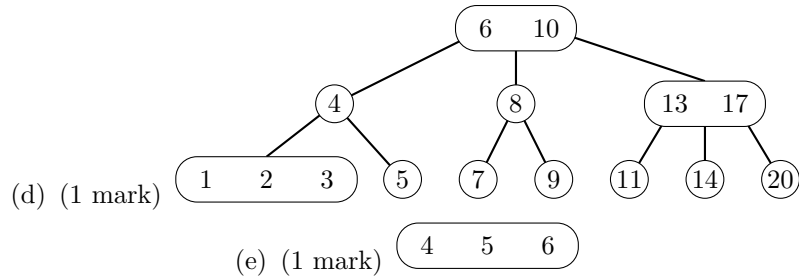6. (5 marks) (Short question.) Consider the 11-vertex weighted graph G below.

What is the weight of a minimum spanning tree of G?

A. 12.

B. 13.

C. 14.

D. G doesn't have a minimum spanning tree.

E. None of the above.

7. (5 marks) (Short question.) Which of the following are valid 2-3-4 trees?

[Implement as jumbled sentence, options "Valid" and "Invalid" for each tree.]

(a) (1 mark)

(b) (1 mark)

(c) (1 mark)

6 10

4

8

13 17

(d) (1 mark) 1 2 3   5   7   9   11 14 20

(e) (1 mark) 4 5 6

8. (5 marks) (Short question.) Consider the unoptimised version of the union-find data structure, in which a sequence of n operations takes $O(n \log n)$ time rather than $O(n\alpha(n))$ time. A user creates a new instance of the structure with MakeUnionFind$(1, \ldots, 100)$, then applies ten Union operations to the resulting data structure. Without knowing the values of those Union operations, what is the **greatest possible** depth of any tree component of the resulting data structure?

   A. 1.

   B. 2.

   C. 3.

   D. 4.

   E. None of the above.

9. (5 marks) (Medium question.) Consider the 2-3-4 tree below:

12

8

16

3 6

10

14

18 20 22

1 2   4 5   7   9   11   13   15   17   19   21   23 24 25
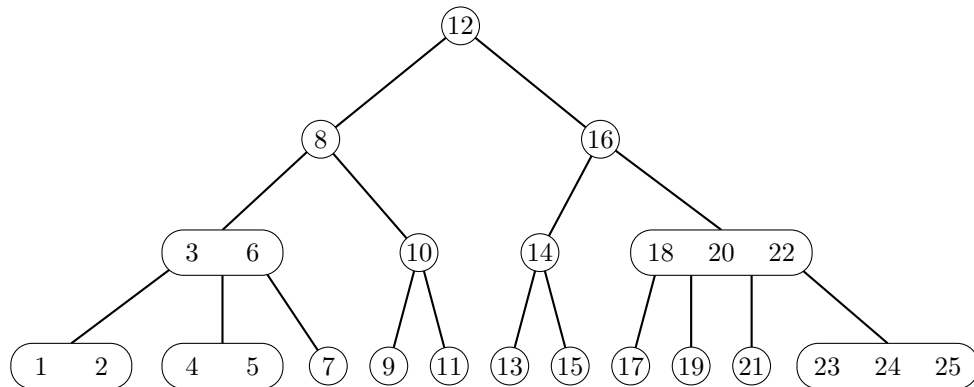
Let T be the result of first inserting the value 30 and then deleting the value 12, assuming that whenever the deletion algorithm has a choice between a "fuse" operation and a "transfer" operation it chooses to transfer. Calculate T.

   (a) (1 mark) What is the depth of T?

   (b) (1 mark) How many children does the root of T have?

   (c) (1 mark) How many 3-nodes does T have?

   (d) (1 mark) How many 4-nodes does T have?

   (e) (1 mark) While inserting 30 and deleting 12, how many fuse, transfer and split operations were performed in total?

10. (5 marks) (Short question.) Consider the following flow network (G, c, s, t) with the following flow f:

s   4/4   a   5/5   b

2/3

1/1   5/5

2/2

2/2

c   3/6   d

4/6

2/2

t
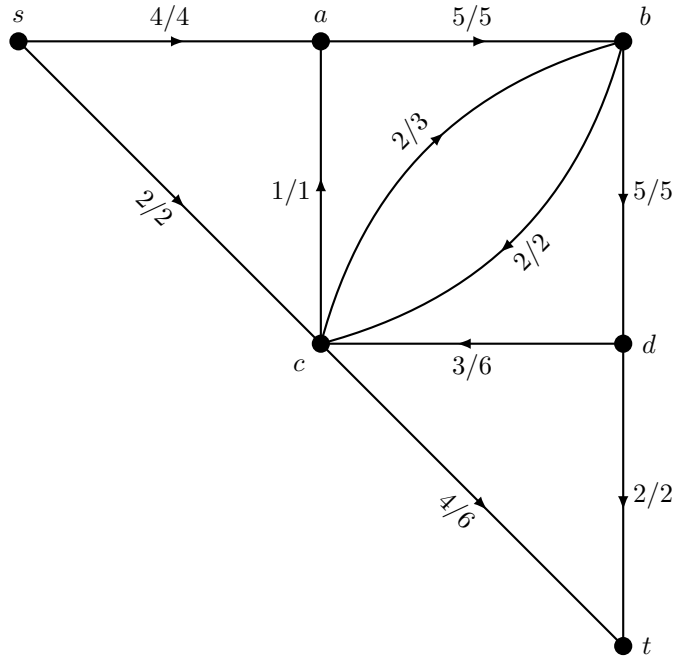
Let $X = \{s, b, d\}$. What is $f^-(X)$?

    A. 7.

    B. 12.

    C. -6.

    D. $f^-(X)$ is not meaningfully defined.

    E. None of the above.

11. (5 marks) (Short question.) Consider the following vertex flow network with source s and sink t:

a   t

1   4

2   1   2

3

s   3   b
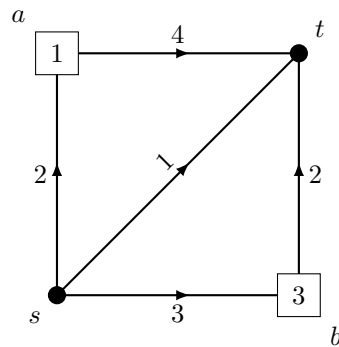
Which of the following is a valid flow f?

    A. $f(s, a) = 1$; $f(s, b) = 3$; $f(s, t) = 1$; $f(a, t) = 1$; $f(b, t) = 2$.

    B. $f(s, a) = 1$; $f(s, b) = 1$; $f(s, t) = 1$; $f(a, t) = 1$; $f(b, t) = 1$.

    C. $f(s, a) = 2$; $f(s, b) = 2$; $f(s, t) = 2$; $f(a, t) = 2$; $f(b, t) = 1$.

    D. $f(s, a) = 1$; $f(s, b) = 3$; $f(s, t) = 1$; $f(a, t) = 1$; $f(b, t) = 3$.

    E. None of the above, or more than one of the above.

12. (5 marks) (Short question.) Consider the graph G shown below, and mark each statement true or false.

b f

a d e h

c g
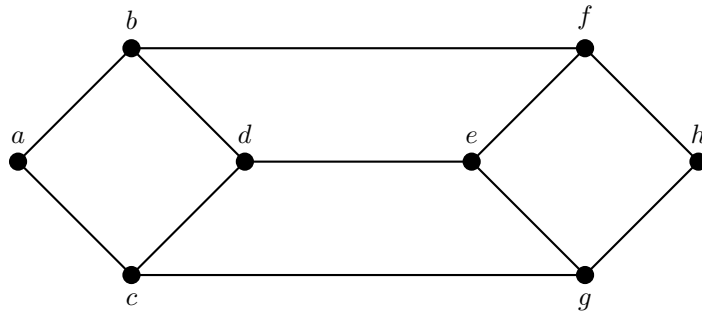
    A. $\{a, d, e, h\}$ is a vertex cover of G.

    B. $\{a, b, c, d, e, f, g, h\}$ is a vertex cover of G.

    C. $\{a, d, e, h\}$ is an independent set of G.

    D. $\{b, c, f, g\}$ is an independent set of G.

    E. If $X \subseteq V(G)$ is an independent set of G, then $V(G) \setminus X$ is a vertex cover of G.

13. (5 marks) (Short question.) Consider an instance of weighted interval scheduling with interval set $\mathcal{R} = [(1, 5), (3, 7), (6, 10), (4, 11), (7, 12), (10, 14), (11, 16), (15, 21), (18, 23)]$ and weight function w given by

$$w(1, 5) = 1, \qquad w(3, 7) = 3, \qquad w(6, 10) = 1, \qquad w(4, 11) = 5, \qquad w(7, 12) = 7,$$
$$w(10, 14) = 6, \qquad w(11, 16) = 4, \qquad w(15, 21) = 4, \qquad w(18, 23) = 6.$$

What is the maximum possible weight of a compatible set?

    A. 13.

    B. 14.

    C. 15.

    D. 16.

    E. None of the above.

14. (5 marks) (Medium question.) Consider the problem of finding a minimum vertex cover on the graph below.
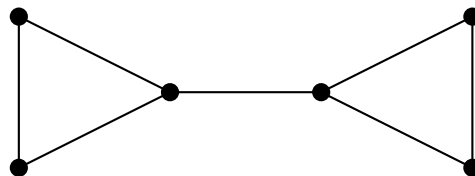
As described in lectures, this problem can be reduced to integer linear programming. Here is one such possible reduction, with some parts missing.

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \rightarrow \text{BLANK subject to:}$$
$$x_1 + x_5 \geq a,$$
$$x_1 + \text{BLANK} \geq a,$$
$$\text{BLANK} + x_2 \geq a,$$
$$\text{BLANK} + x_6 \geq a,$$
$$x_6 + x_4 \geq a,$$
$$x_5 + x_4 \geq a,$$
$$x_3 + x_5 \geq a,$$
$$0 \leq x_1, \ldots, x_6 \leq a,$$
$$x_1, \ldots, x_6 \in \mathbb{N},$$
$$a = \text{BLANK}.$$

Fill in the blanks to make a valid reduction. Note that the correspondence between variables in the linear program and vertices in the graph is not given. This is not a mistake in the question; you should be able to fill in the blanks without it. (Any valid solution will get full marks.)

Options for blanks: min, max, x_1, x_2, x_3, x_4, x_5, x_6, 0, 1 and 2.

15. (10 marks) (Short question.) Consider a maze, as shown below.
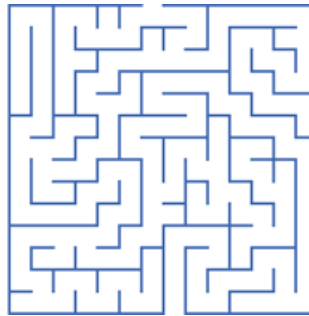
Explain very briefly how you would find the shortest path from the entrance to the exit in $O(n \log n)$ time, where n is the number of junctions in the maze. You may assume that as part of the input, you have access to:

- A list J of all junctions in the maze plus the entrance and exit;
- A list C of all corridors directly connecting two junctions, along with their lengths.

You may assume $|C| = O(n)$; this is true for any 2D maze. You may use any algorithm covered in the course without explaining its implementation, and you do not need to prove your algorithm works.

16. (10 marks) In a graph $G = (V, E)$, we say a vertex $v \in V$ is *critical* if G is connected and $G - v$ has exactly two components, i.e. if deleting v and all its edges from G splits G into two parts.

   (a) (5 marks) (Short question.) For $n \geq 3$, what is the **least** number of edges an n-vertex graph can have while still having a critical vertex, and why?

   (b) (5 marks) (Medium question.) For $n \geq 3$, assuming n is odd, what is the **greatest** number of edges an n-vertex graph can have while still having a critical vertex, and why? You may wish to use the following algebraic fact, which you do not need to prove: For all $a, b, t > 0$ satisfying $a + b \leq t$ we have $a^2 + b^2 \leq t^2/2$, i.e. $a^2 + b^2$ is maximised when $a = b = t/2$.

17. (10 marks) (a) (5 marks) (Short question.) How large is a maximum matching for the graph below? Briefly explain your reasoning.
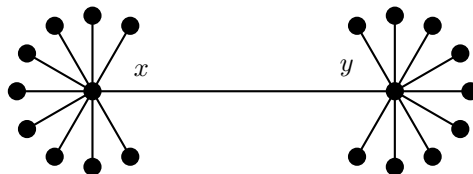
(b) (5 marks) (Medium question.) Give an example of a graph with no perfect matching in which every vertex has degree 16, and briefly explain why it has no perfect matching. Your graph does not have to be bipartite, and you do not have to draw it.

18. (10 marks) (Medium question.) After making some questionable life choices, you are developing marketing materials for a company selling stock trading advice to the general public. To illustrate the power of good financial advice, you want to know how much money someone could theoretically have made on Gamestop shares if they had initially invested in one share and then sold and re-bought a limited number of times with perfect foreknowledge. ("With just five smart trades at the right moments, you could have earned over five hundred squillion pounds! Trade the smart way. Trade with StonksCo!")

You have access to a list $p_1, \ldots, p_n$ of Gamestop's share prices at the start of each day of historical trading; you may assume the price stays constant throughout the day and is unaffected by your theoretical trades, and that you can buy fractions of shares. You are also given a number k; this is the maximum number of trades you are allowed to make after buying the initial share on day 1. The only possible trades you need to consider are selling all your shares (when the price is higher) or reinvesting all your money into buying new shares (when the price is lower). Your goal is to find the maximum possible profit, in either stocks or cash. Fill in the blanks of the following dynamic programming algorithm to solve the problem.

Here, $H[i, j]$ will contain the maximum amount of money available from owning a single share at the start of day j with i trades remaining. You should infer the meaning of $N[i, j]$ from the algorithm.

```
1  begin
2      Initialise H and N to empty two-dimensional arrays indexed by {0, ..., k} × {1, ..., n}.
3      Initialise H[i, n] = pₙ for all i ≤ k.
4      Initialise N[i, n] = pₙ for all i ≤ k.
5      Initialise H[0, j] = pₙ for all j ≤ n.
6      Initialise N[0, j] = pⱼ for all j ≤ n.
7      for i = 0 to k do
8          for j = n − 1 to 1 do
9              Let H[i, j] = max{BLANK, max{N[i − 1, k]: k ≥ j + 1}}.
10             Let N[i, j] = max{BLANK, max{BLANK/BLANK · BLANK[i − 1, k]: k ≥ j + 1}}.
11     Return H[1, k] − p₁.
```

Options for blanks: p_1, p_i, p_j, p_k, p_n, H, N.

Section 2

19. (10 marks) (Medium question.) Consider the following greedy algorithm `heuristicIS`, intended to find a maximum independent set in an arbitrary graph in polynomial time.

Input : A graph G with vertex set $\{1, 2, \ldots, n\}$.
Output: An independent set X of G which we hope is maximum.
1 begin
  2   Let $X \leftarrow \emptyset$.
  3   **while** $V(G) \neq \emptyset$ **do**
  4   Let $\delta \leftarrow \min\{d(v) \colon v \in V\}$.
  5   Let x be the lowest-numbered vertex with $d(x) = \delta$.
  6 $X \leftarrow X \cup \{x\}$.
  7 $G \leftarrow G - x - N(x)$, the graph with vertex set $V(G) \setminus (N(x) \cup \{x\})$ and edge set
       $\{e \in E(G) \colon e \cap (N(x) \cup \{x\}) = \emptyset\}$.
  8   Return the value of X.

(a) (1 mark) Why, without looking at the implementation of `heuristicIS`, can we be almost certain that it doesn't actually return a maximum independent set for all input graphs?

(b) (4 marks) Give an example of a graph G and a *key vertex* $v \in V(G)$ such that every vertex in G has degree at least two, and such that every maximum independent set of G contains v.

(c) (5 marks) Using your answer to part (b) or otherwise, give an example of a graph on which `heuristicIS` fails to return a maximum independent set. Explain why your example works.

20. (10 marks) (a) (5 marks) (Long question.) Let $G = (V, E)$ be a connected graph with edge weights given by $w \colon E \to \mathbb{R}$. You may assume that every edge gets a different weight. Let C be a cycle in G, and let e be the highest-weight edge in C. Prove that no minimum spanning tree of G contains e. You may assume any consequence of the Fundamental Lemma of Trees without proof as long as you state it clearly.

(b) (5 marks) (Medium question.) Using the result stated in part (a) or otherwise, give an algorithm which, given an n-vertex connected graph $G = (V, E)$ in adjacency list format with $|E| = n + 50$, outputs a minimum spanning tree in $O(n)$ time. Briefly explain why your algorithm works and why it runs in $O(n)$ time.

21. (10 marks) (Long question.) As in a previous question, consider a maze. This time, however, every junction in the maze contains a button, and some corridors are blocked by gates. On reaching a button, you have the option of pressing it as many times as you like; this will close some gates, open others, and toggle others (i.e. they open if they are closed and close if they are opened). See the picture below for an example.
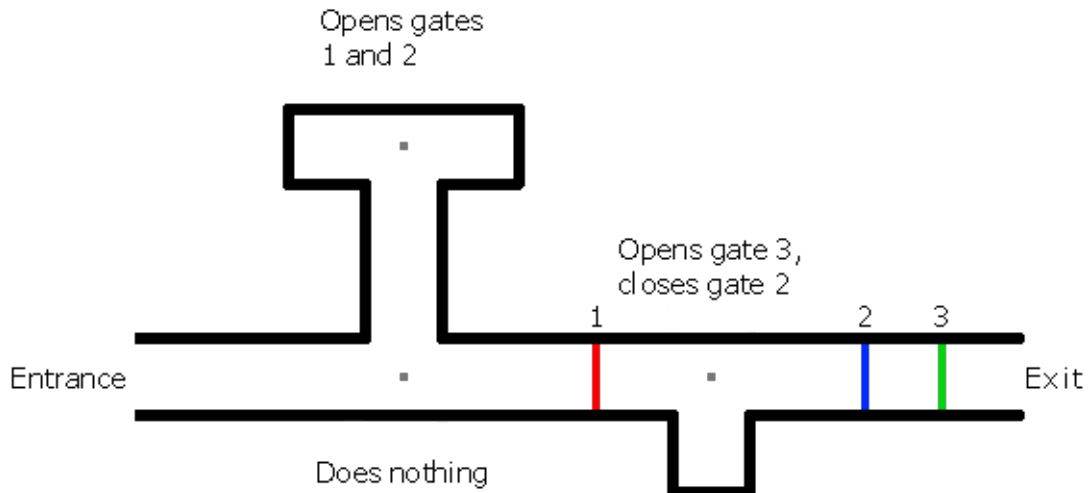
Briefly describe an algorithm to find the shortest route from the entrance to the exit in $O(2^g n(g+\log n))$ time, pressing as many buttons as you like, where n is the number of junctions in the maze and g is the number of gates. You may assume the input is given as:

- A list J of all junctions in the maze plus the entrance and exit and the behaviour of their buttons;

- A list C of all corridors directly connecting two junctions, along with their lengths;

- A list G of all gates, along with their initial states (open or closed) and which corridors they are contained in.

You may assume $|C| = O(n)$; this is true for any valid 2D maze. You may use any algorithm covered in the course without explaining its implementation, and you do not need to prove your algorithm works. (**Hint:** Try reducing the problem to finding a single shortest path between two vertices in a graph.)

22. (10 marks) **Important:** The two parts of this question are completely independent — they have been put together for technical Blackboard reasons. Doing part a) will not help you with part b).

(a) (5 marks) (Medium question.) You are attempting to form a committee of people to represent Computer Science in the upcoming faculty restructure. A regrettably large number of stakeholders have ideas about what this committee should look like: Denoting the stakeholders by $S_1, \ldots, S_n$, each stakeholder S_i has one list S_i^+ of people they would like to be on the committee, and another list S_i^- of people they would like not to be on the committee. (Either list may be empty.)

You quickly realise that it is impossible to satisfy everyone's preferences at once, so you decide to try for something easier: you are trying to grant every person at least one of their requests. Thus for every stakeholder S_i, either you have added a person in S_i^+ to the committee, or you have *not* added a person in S_i^- to the committee, or both. For example, if everyone requested that John Lapinskas be added to the committee, then any committee containing John would be a valid committee. Sketch a proof that even so, deciding whether or not a valid committee exists given $S_1^+, \ldots, S_n^+$ and $S_1^-, \ldots, S_n^-$ is an NP-complete problem.

(b) (5 marks) (Long question.) Let $G = (V, E)$ be an n-vertex connected graph. For each $v \in V$, let T_v be a BFS tree of G rooted at v. Let $d_G(x, y)$ denote the distance between x and y in G, and let $d_v(x, y)$ denote the distance between x and y in T_v. Let

$$\sigma(G) = \sum_{(x,y) \in V} d_G(x, y), \qquad \sigma_v = \sum_{(x,y) \in V} d_v(x, y),$$

where both sums are over all **ordered** pairs x, y (so (x, y) and (y, x) are two different terms). Prove that $\sum_{v \in V} \sigma(T_v) \leq 2(n-1)\sigma(G)$.