Feedback About 2021's Exam in MATH20902: Discrete Mathematics

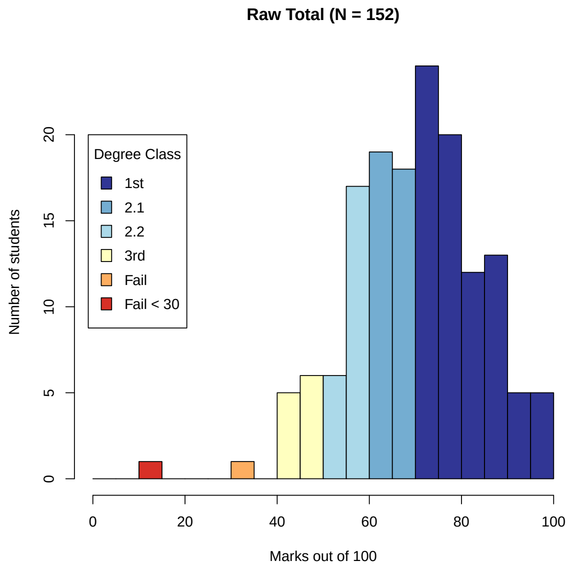


Figure 1: *A histogram of the sum of the three components of the assessment. These are "raw" marks in the sense that they have neither been moderated (that is, scaled) by the departmental Exam Board, nor adjusted to account for any Mitigating Circumstances, DASS allowances or late penalties.*

General remarks

- The assessment this year had three components: an online courswork test (20%) on which most students did very well; an on online component to the end-of-term exam (30%) where, again, most students did well and a take-home, written component (50%). Figure 2 on the next page shows histograms of the marks in these three components.
- Many people did very well on the course: here is a summary of the raw marks, by degree class.

Result:	First	2(i)	2(ii)	Third	Fail
Range:	70–100	60–69	50–59	40–49	0–39
Number of students:	79	37	23	11	2
Fraction of students:	52.0%	24.3%	15.1%	7.2%	1.3 %

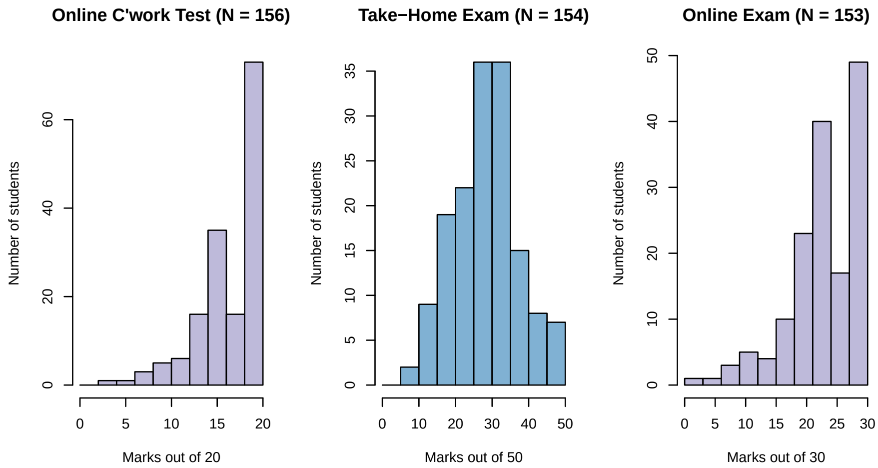


Figure 2: *Histograms of the results from the three components of the assessment: at left, the online coursework test; in the middle, the written, take-home of the exam; and, at right, the online component of the exam. Note that the vertical scale and the total number of marks available varies from component to component.*

The take-home component of the exam

I was interested in how people managed their time and so made Figure 3, which is a histogram of the time, in minutes, at which students submitted the take-home component of the exam. Most people did so in the last 15–20 minutes before the noon deadline.

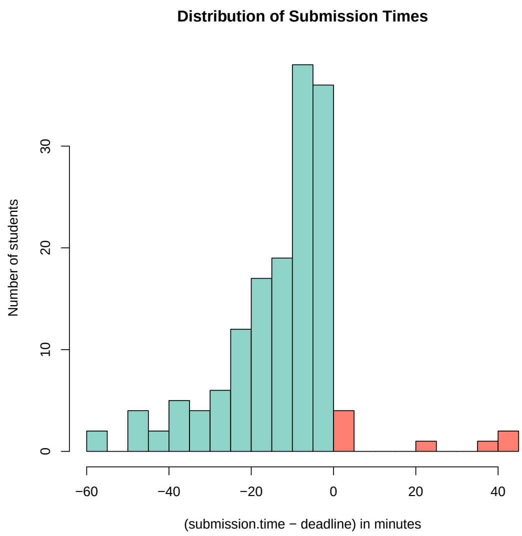


Figure 3: *The distribution of the times histogram of the times, in minutes, at which students submitted the take-home component of the exam. I've shifted the origin so that the deadline falls at* $t = 0$ *and thus negative times lie within the submission window. The handful of counts representing late submission may correspond to students who are allowed extra time, but I have no easy way to check this.*

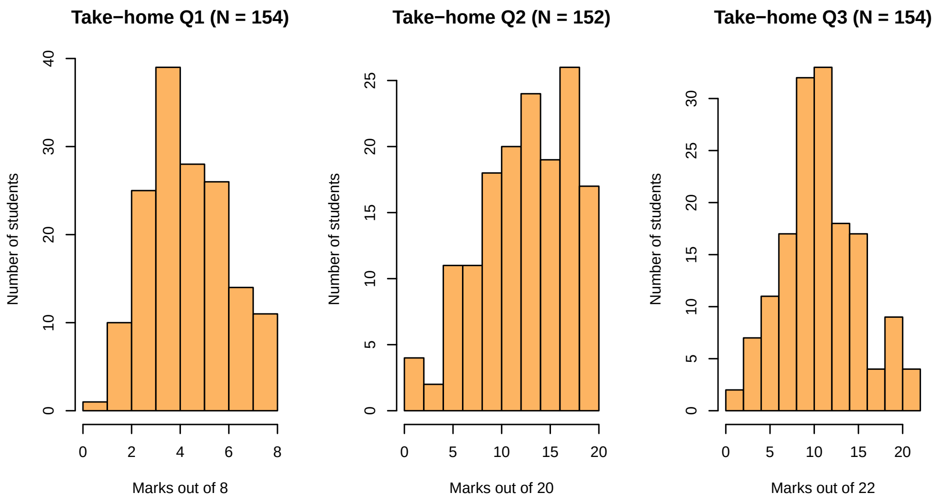


Figure 4: *Histograms of marks for the individual questions from the take-home component of the exam: note that the vertical scale and the total number of marks available varies from question to question.*

Q1 People did reasonably well on this question: the average was around 4.8/8, with 11 perfect scores.

- The main way people lost marks (a maximum of 3) was by failing to explain the construction of the graph. I wanted to see:
  - a remark along the lines of "The vertices correspond the language exams . . . ";
  - an explanation that edges connect pairs of exams taken by the same student;
  - an explanation of how the graph relates to the scheduling problem. This could be as brief as: "If we then colour the graph, the only exams that receive the same colour will be those that can be held simultaneously and so the chromatic number of the graph is also the minimal number of exam sessions."

  In what follows, I'll refer to the graph as G and to the vertices with lowercase letters, so that the vertices for Arabic, Bengali and Czech are, respectively, a, b and c.
- Many, perhaps most, students simply found a six-colouring and then declared that this proved $\chi(G) = 6$. This is incorrect: an example of a six-colouring establishes only that $\chi(G) \leq 6$.
- Those who did try to give a complete proof typically used one of the following strategies:
  - They claimed that G contains a subgraph isomorphic to K_6, which would establish that $\chi(G) \geq 6$. Unfortunately, this just isn't true, so this sort of argument won few marks.
  - Others noted that G contains many subgraphs isomorphic to K_5, so that $\chi(G) \geq 5$. Some the went on to probe, typically by contradiction, that $\chi(G) \neq 5$. This was the most successful strategy.
  - Quite a few students noticed that the vertices f, g and h are adjacent to each other and to all others and that the five remaining vertices lie on a cycle of odd length. These observations allow one to prove two things:
    * G has a six-colouring;
    * G can't have a five-colouring.

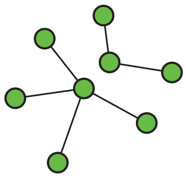
    If this logic was laid out clearly, I awarded full credit, but if the points above were muddled together, I deducted a mark or two.

Q2 Here too, people did reasonably well: the average was around 12/20, with 8 perfect scores and 54 students who earned scores of 16 or more.

The question tests material from the videos from Week 9, especially those about the chapter in the notes entitled *Tropical Arithmetic and Shortest Paths*. Parts (a)–(c) depended on a theorem that says that if A is the adjacency matrix of a graph, then $A^{\ell}_{i,j}$ is the number of walks of length ℓ from v_i to v_j.

(a) Both propositions here are true.

- The first one is an immediate consequence of the theorem mentioned above and the definition of a connected graph, and most students realised this. One surprisingly common mistake was the claim that if a graph $G(V, E)$ is not connected, then there must be some $v \in V$ such that $\deg(v) = 0$. A little thought shows this can't really be true, and the graph below—which is not connected, yet has no isolated vertices—provides a counterexample.



- The proof of the second proposition turns on the observation that if, as usually, we write the vertex set of the bipartite graph $G(V, E)$ as the union of two non-empty, disjoint sets V_1 and V_2, then a vertex sequence $(v_0, v_1, \dots, v_\ell)$ that specifies a walk must alternate between V_1 and V_2. That is, if $v_0 \in V_1$, then $v_j \in V_1$ for all even j and $v_j \in V_2$ for all odd, j. In light of the theorem mentioned above, this means that if k is odd, $A^k_{1,1} = 0$, while if k is even and $(v_i, v_j) \in E$, it must be true that $A^k_{i,j} = 0$. Thus A^k always has at least one entry equal to zero, for any value of k.

The most common mistake here was the somewhat baffling claim that K_2 or $K_{1,1}$ provide a counterexample to the second proposition. This is an odd thing to say, as both these graphs have

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

and so

$$A^k = \begin{cases} A & \text{If } k \text{ is odd} \\ I_2 & \text{If } k \text{ is even} \end{cases}$$

where I_2 is the 2×2 identity matrix. Either way—k odd or even—A^k always has two entries equal to zero.

(b) The idea here is that $A^{31}_{3,4}$ counts walks of length 31 from v_3 to v_4. Brief study of the graph shows that any such walk must consist of:

- $p \geq 0$ trips around the three-cycle (v_3, v_1, v_2, v_3), followed by ...
- a step from v_3 to v_4 and, finally, ...
- $q \geq 0$ trips around the two-cycle (v_4, v_5, v_4).

The problem thus reduces to finding all integer solutions to

$$3p + 1 + 2q = 31 \tag{1}$$

where $p \geq 0$ and $q \geq 0$. There are six of them.

Among students who received only partial credit there were two common mistakes:

- Counting only some of the solutions to Eqn. (1), typically the cases $p = 10, q = 0$ and $p = 0, q = 15$.

- Simply writing down the matrix A^{31}, where A is the adjacency matrix of the graph in the exam, and then reading off $A^{31}_{3,4}$. If the student explained how they computed A^{31}, I was more generous than if it just appeared out of thin air (or perhaps out of *Wolfram Alpha*).

(c) Here again, the idea is to count walks, this time in K_n and of the form (v_1, v_a, v_b, v_1). It's not hard to see that there are $n-1$ choices for v_a (which must differ from v_1) and $n-2$ choices for v_b (which must difer from both v_1 and $v_a \neq v_1$), so that $A^k_{1,1} = (n-1)(n-2)$.

I saw:

- several examples of an ingenious argument that starts with the observation that if A is the adjacency matrix of K_n, then $A = J_n - I_n$, where J_n is the $n \times n$ matrix entirely filled with 1's and I_n is the $n \times n$ identity matrix. Then, because J_n and I_n commute, we have
$$\begin{aligned} A^3 &= (J_n - I_n)^3 \\ &= J_n^3 - 3J_n^2 I_n + 3J_n I_n^2 - I_n^3 \\ &= J_n^3 - 3J_n^2 + 3J_n - I_n \end{aligned}$$
which makes it easy to see that $A^3_{1,1} = n^2 - 3n + 2$.
- many baffling computations that began with, for example, the assertion that the adjacency matrix of K_5 is
$$A = J_5 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} ?!?$$
even though K_5 has no loops.

(d) I had hoped that students would recognise that the matrix W in this question is an example of the weight matrix defined in the section of the notes entitled *A tropical version of Bellman's equations* and so realise that the smallest entry is $W^{\otimes 4}$ is the weight of the lowest-weight walk of length 4 in the graph. A few people did realise this, but the vast majority of students tried to calculate $W^{\otimes 4}$ directly. This often worked, but also often led to mistakes. Surprisingly few students used repeated-squaring, so did three matrix multiplications rather than two.

Q3 This was the hardest problem and also the last, so that people may have been running short of time when they reached it (though, of course, one needn't attempt the problems in the order they're printed). Nevertheless, students did reasonably well and the average was just over 10.6/22, with 10 students getting a score of 20 or more.

(a) I meant this part to be easy and students seem to have found it so: the vast majority got 4 or 5 marks. Where people lost a marks it was usually for one of the following reasons:

- They didn't provide any justification for the claim that $m = d\,2^{d-1}$: I wanted to see a small calculation based on the Handshaking Lemma.
- They didn't justify the claim that $g = 4$ when $d > 1$: here I wanted to see some explanation about why I_d never contains a 3-cycle. It was enough to say "I_d is bipartite for all d, so any cycles must have even length."

Quite a lot of students seemed to feel obliged to assign a girth to I_1, which doesn't contain any cycles. I saw both $g = 0$ and $g = \infty$, but the correct statement is that the girth is not defined for an acyclic graph. I didn't deduct any marks for these mistakes. Finally, a handful students seem to have gotten in a muddle about the definition girth and so tried to use the inequality in Eqn. (2) to compute g: this was never successful.

(b) The answers to the main questions here—for which d is I_d Eulerian and/or Hamiltonian—appeared in the problem sets and so most people got at least two of the five marks available here. For full credit I wanted to see:

- In the part about Eulerian tours, a mention that I_d is *connected*. This is because the key theorem here, which says that a graph is Eulerian iff all its vertices have even degree, only applies to connected graphs.
- In the part about Hamiltonian tours, some discussion of the recursive algorithm that constructs a 2^d-cycle in I_d.

Quite a few people seem to have just copied the relevant passages out of the solutions to the problem sets. This led them to write things such as

> Given that I_d has exactly 2^d vertices, the same construction yields a Hamiltonian cycle: it's just that we didn't know to call it that earlier in the term. That is, I_d is Hamiltonian for $d \geq 2$.

without ever saying how the construction works: this cost a couple of marks.

Finally, some students tried to use Dirac's theorem (or Ore's or the Bondy-Chvátal theorem) to check whether I_d is Hamiltonian. This leads to the correct result for $d = 2$, but is inconclusive for $d > 2$. Many of these people went on to make a logic mistake: theorems such as Dirac's, which say "high degree $\Rightarrow$ Hamiltonian", can only be used to prove that a graph *is* Hamiltonian, not that it *isn't*.

(c) Almost everyone realised that one should use bounds on the number of edges here, but lots of people tried the weaker $m \leq 3n - 6$, which allows you to prove that I_d can't be planar for $d \geq 6$, rather than the stronger

$$m \leq \frac{g(n-2)}{g-2}, \tag{2}$$

which establishes that I_d is nonplanar for $d \geq 4$. Other notes about this part include:

- Surprisingly few people remarked that I_1, I_2 and I_3 must be planar as the exam paper includes planar diagrams for them.
- Many people made the clear and useful observation that as I_4 is a subgraph of all I_d with $d \geq 4$, so if you can prove that I_4 is nonplanar, then you've also proven that all the others are nonplanar too.
- Lots of people used the bounds above to establish that most cube graphs are nonplanar, but went on to claim that this proves that I_d is planar for $d < 4$. This is a mistake in logic: bounds such as those above say something like "planar graphs can't have many edges" and so cannot be used to prove that a graph *is* planar, only that it *isn't*.
- Finally, a few students tried to make arguments based on Euler's formula: $f = m - n + 2$. It's hard to see how this could work, as faces are a feature of a planar diagram and so aren't even well-defined unless one knows the graph in question is planar. It's possible that these arguments could be reframed as proofs by contradiction: "Assume for contradiction that I_4 is planar. Then it must have ... some impossibly large number of faces bounded by 4-cycles", but none of the attempts I saw were as clear as this and so they didn't earn many marks.

(d) I thought people would find this part challenging and most did, but a respectable number of students found suitable examples: there were 11 perfect scores. I awarded a few marks for sensible applications of Kuratowski's theorem, but a lot of people went on to claim something like "I_d can't contain a subgraph homeomorphic to K_5". This isn't true and the arguments offered in support, which usually had something to do with the degrees of vertices being too high, don't work because we're looking for *subgraphs* homeomorphic to K_5 or $K_{3,3}$. This means we're free to delete edges or vertices as needed. In fact, all nonplanar cube graphs contain I_4 as a subgraph and so, as Figure 5 illustrates, all contain subgraphs homeomorphic to *both* $K_{3,3}$ and K_5.

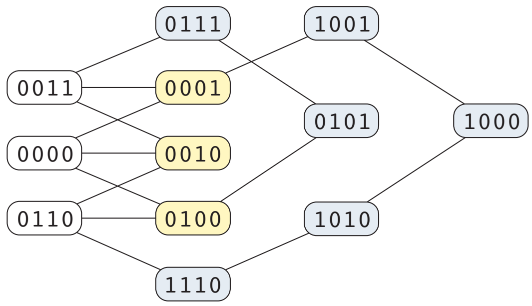


Figure 5: *At left, a subgraph of the cube graph I_4 that's homeomorphic to $K_{3,3}$. The vertices corresponding to the two parts of $K_{3,3}$'s vertex set are shown in white and yellow, while the vertices inserted by subdivision are pale blue. At right is a similar diagram showing that I_4 contains a subgraph homeomorphic to K_5.*

The online component of the exam

Each student received an individualised test consisting of three questions—one each about strongly connected components, critical path analysis and counting spregs—in which each question was drawn uniformly at random from a large pool of similar questions. There were 150 questions per pool, which means there were potentially $150^3 = 3375000$ distinct exams.

It's the first time I have prepared an online component to the final exam and I have yet to understand fully the analyses that Blackboard generates, but I have provided a few remarks below.

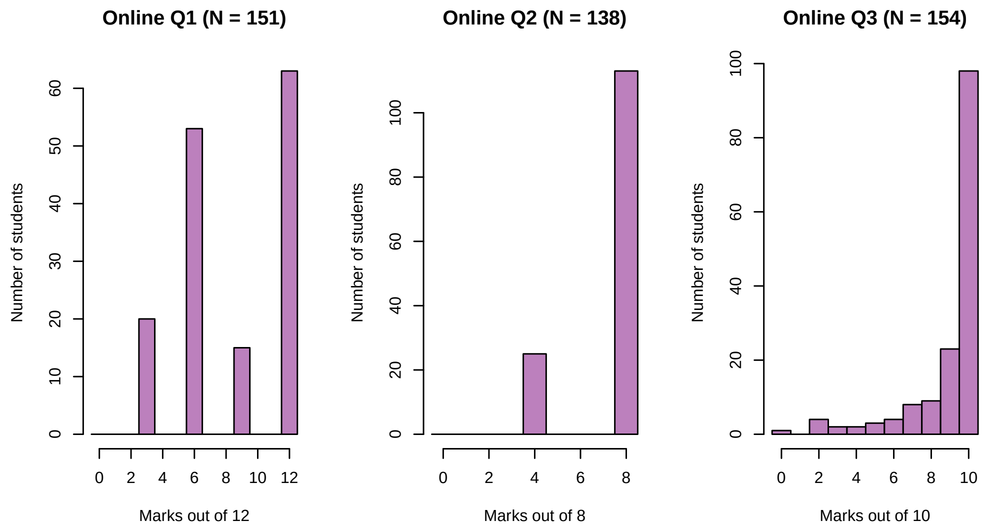


Figure 6: *Histograms of marks for the individual questions from the online component of the exam: note that the vertical scale and the total number of marks available varies from question to question.*

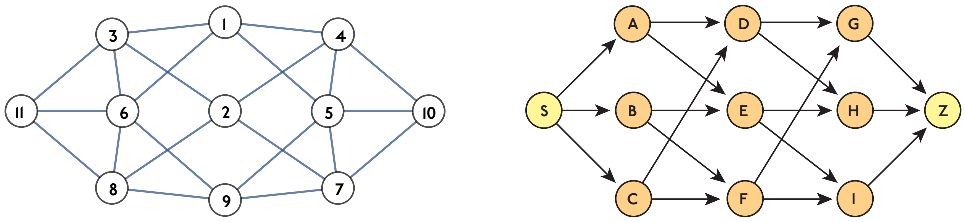


Figure 7: *Graphs underpinning the questions about connected components (left) and critical path analysis (right).*

Connected Components The average for this question was 8.2/12 and 63 students got full marks. The marking scheme was set up to award partial credit and to accept lists of vertices in arbitrary order. Thus if the answer was "2 7 9", the lists "2 9 7" and "9 2 7" would also be correct. All 150 questions in the pool involve a digraph derived from the one at left in Figure 7 by choosing directions for the edges at random. All the digraphs appearing in the exam:

- are unique in the sense that none is isomorphic to any of the others;
- have at most 8 strongly connected components;
- have at least one component with three members.

The diagrams that appeared in the exam were drawn automatically by *Mathematica* and most look like directed versions of the one in Figure 7, but four or five were rather harder to interpret. For those cases, the Teaching and Learning Office and I reviewed each student's answers carefully to make ensure that no disadvantage occurred.

Counting Spregs The average for this question was 6.5/8 and 113 students got full marks. Here too, the automated marking was arranged so as to award partial credit. All 150 questions in the pool involve a digraph that:

- is unique in the sense that it is not isomorphic to any of the others;
- has 6 vertices and 12 directed edges;
- contains at least one cycle of length 2, 3, or 4;
- contains at least one spreg.

Here too, a handful of the automatically generated diagrams were hard to interpret and, for students who got those questions, their answers were reviewed carefully to ensure fairness.

Critical Path Analysis The average for this question was 9/10 and 113 students got either 9.5 or 10 marks. The marking scheme was designed to award partial credit and to accept many representations of the critical path. In particular, if the critical path is S-C-F-G-Z then any of the following would (along with many other variants) receive full credit:

"S C F G Z", "C, F, G, Z", "C F G", "CFG", "s C f G z" and "s-c-f-g-z".

All 150 questions involve the same underlying graph—shown at right in Figure 7—but have different durations for the tasks and thus different edge weights, critical paths and minimal times to completion. All the critical path problems:

- are unique in the sense that they have distinct sets of durations for the tasks;
- involve task durations τ_v that are integers in the range $5 \leq \tau_v \leq 15$;
- have just one critical path.