

MATH 3012 Final Exam, May 3, 2013, WTT

Note. The questions posed here appeared on the Spring 2013 Final Exam, although in some instances, the format of the question has been altered. Also, answers provided here are more comprehensive than students are expected to provide in answering the given questions, and in some cases, extensive discussion of background issues has been provided. This more comprehensive version is intended to serve as a study guide, and should be especially meaningful to students enrolling in Math 3012 in future semesters.

1. Consider the 62-element set X consisting of the twenty-six letters (case sensitive) of the English alphabet and the ten digits $\{0, 1, 2, \dots, 9\}$.

a. How many strings of length 15 can be formed if repetition of symbols is permitted?

Standard string problem. The number of strings of length n from an m -letter alphabet is m^n , so here the answer is 62^{15} .

Note. In our class, leaving an answer in the form 62^{15} is allowed since one can readily carry out the arithmetic to find the precise answer. Here Maple reports that

$$62^{15} = 768909704948766668552634368$$

For the remainder of these notes, I have used Maple to make complicated computations, and I strongly encourage you to experiment with a tool, such as Maple or Mathematica, to see how fast arithmetic calculations can be made and to learn where even the best of tools run out of gas.

b. How many strings of length 15 can be formed if repetition of symbols is *not* permitted?

Standard permutation problem. The number of permutations of length n from an m -letter alphabet is $P(m, n)$, where $P(m, n) = m!/(m - n)!$ when $m \geq n$ and $P(m, n) = 0$ when $m < n$. Of course, when making this calculation, we take advantage of the fact that all factors in $(m - n)!$ cancel. So in our specific problem, the answer is:

$$P(62, 15) = 121682695942193436610560000$$

c. How many strings of length 15 can be formed using exactly four A 's, two a 's, seven 3's and two 5's?

Standard "Mississippi" problem. Here the answer is

$$\binom{15}{4, 2, 7, 2} \text{ or } \frac{15!}{4!2!7!2!} = 2702700.$$

d. How many strings of length 15 can be formed if exactly seven characters are capital letters, exactly four characters are 6's and the remaining four characters are digits? Here, repetition is allowed.

The answer is

$$\binom{15}{7} 26^7 \binom{8}{4} 9^4 = 23737231472085331200.$$

Note that exactly four of the characters are 6's. This means that the other digits are drawn from a set of size 9.

- e. How many binary relations are there on X ?

There are n^2 ordered pairs that can be formed from a set X of size n , and a binary relation on X is just a subset of $X \times X$. So the number of binary relations on X is 2^{n^2} . In our problem, the answer becomes 2^{6^2} .

Note that I didn't report on the precise answer, as provided by Maple, even though Maple was able to calculate 2^{6^2} in the blink of an eye. However, the answer would require a full page of space to print. Check this out for yourself!

- f. How many equivalence relations are there on X with class sizes 15, 15, 6, 6, 6, 6 and 2?

The answer is

$$\frac{\binom{15}{4,2,7,2}}{2!5!1!} \quad \text{or} \quad \frac{15!}{4!2!7!2!5!1!} = 98905131133688880.$$

The additional terms in the demonator stem from the ways of permuting the classes of a specified size.

Here are some more general remarks about the number of relations of a particular type.

When X is an n -element set:

1. The number of symmetric binary relations on X is $2^{\binom{n}{2}}$. Note that the relation must include each pair of the form (x, x) . When $x \neq y$, either the relation contains both (x, y) and (y, x) or it contains neither of these two pairs. So for each 2-element subset of X , there are two choices. Note that this is also the number of labelled graphs on n vertices.
2. The number of anti-symmetric binary relations on X is $2^n 3^{\binom{n}{2}}$. For each pair (x, x) , we get to decide whether it belongs to the relation. When $x \neq y$, we have three choices: include (x, y) but not (y, x) , include (y, x) but not (x, y) or exclude both (x, y) and (y, x) .
3. The number of reflexive, anti-symmetric binary relations on X is $3^{\binom{n}{2}}$. For each pair (x, x) , we must put (x, x) in the relation, and when $x \neq y$, we have the same three choices as above.

However, you should note that once the transitivity condition is added, such nice formulas can no longer be found. So it is quite challenging to find the number $p(n)$ of partial orders on an n -element set. Here is a table of values that I found on the web (search for the *On-Line Encyclopedia of Integer Sequences*).

$$\begin{aligned}
p(1) &= 1 \\
p(2) &= 3 \\
p(3) &= 19 \\
p(4) &= 219 \\
p(5) &= 4231 \\
p(6) &= 130023 \\
p(7) &= 6129859 \\
p(8) &= 431723379 \\
p(9) &= 44511042511 \\
p(10) &= 6611065248783 \\
p(11) &= 1396281677105899 \\
p(12) &= 414864951055853499 \\
p(13) &= 171850728381587059351 \\
p(14) &= 98484324257128207032183 \\
p(15) &= 77567171020440688353049939 \\
p(16) &= 83480529785490157813844256579 \\
p(17) &= 122152541250295322862941281269151 \\
p(18) &= 241939392597201176602897820148085023
\end{aligned}$$

I can check the values of $p(n)$ for $n \leq 4$ by hand, but verifying that $p(4) = 219$ takes me about five minutes or so. I would scream bad things before attempting to find $p(5)$ by hand. Not surprisingly, finding these values is a modest programming challenge, say for $n \leq 7$, but for larger values, it is a real computational challenge, and finding the last several values in this table required considerable programming skill, an array of very fast computers and *extensive* computing time.

2. How many integer valued solutions are there to the equation $x_1 + x_2 + x_3 + x_4 + x_5 = 62$ when:
- a. $x_i > 0$ for $i = 1, 2, 3, 4, 5$.

Starting point for these kinds of problems—reformulate as choosing gaps. In general, the number of ways to distribute n identical object into m distinguishable bins, in such a way that all bins are non-empty, is $\binom{n-1}{m-1}$. Of course, this requires $n \geq m$. Here the answer is $\binom{61}{4} = 521855$.

- b. $x_i \geq 0$ for $i = 1, 2, 3, 4, 5$.

Add one artificial object for each bin. Then choose gaps and reduce the number of objects in a bin by one. So here the answer is $\binom{66}{4} = 720720$.

- c. $x_i > 0$ for $i = 1, 2, 3, 5$ and $x_4 > 9$.

Give 9 objects to Bin 4 in advance. Then distribute the remaining 53 objects. So answer is $\binom{52}{4} = 270725$.

- d. $x_i > 0$ for $i = 1, 2, 3, 4, 5$ and $x_4 \leq 9$.

Just the difference between two previous solved problems. Here the answer is $\binom{61}{4} - \binom{52}{4} = 251130$.

- 3 a.** Use the Euclidean algorithm to find $d = \gcd(252, 1320)$.

Long division reports that

$$1320 = 5 \cdot 252 + 60$$

$$252 = 4 \cdot 60 + 12$$

$$60 = 5 \cdot 12 + 0$$

The last non-zero remainder is the greatest common divisor. So here, $d = 12$.

- b.** Use your work in the first part of this problem to find integers a and b so that $d = 252a + 1320b$.

We rewrite the work previously done as:

$$12 = 1 \cdot 252 - 4 \cdot 60$$

$$60 = 1 \cdot 1320 - 5 \cdot 252$$

Substituting we obtain

$$12 = 1 \cdot 252 - 4(1 \cdot 1320 - 5 \cdot 252) = 21 \cdot 252 - 4 \cdot 1320.$$

This shows that we may take $a = 21$ and $b = -4$. Note that there are actually infinitely many solutions to this problem. Here is the general form $a = 21 + 1320k$, $b = -4 - 252k$, where k is an integer.

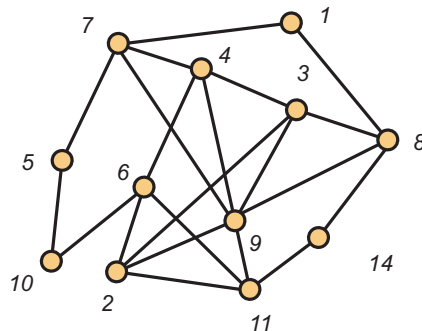
- 4.** For a positive integer n , let s_n count the number of ternary sequences which do not have three consecutive 2's. Note that $s_1 = 3$, $s_2 = 9$, $s_3 = 26$ and $s_4 = 76$. For $n \geq 5$, develop a recurrence for s_n and use it to find s_6 .

Let $n \geq 5$ and consider a good sequence of length n . If it ends in a 0 or a 1, then the first $n - 1$ characters represent any admissible sequence of length $n - 1$. However, if the last character is a 2, then *some* of the admissible strings of length $n - 1$ cannot be used. These strings have the last two characters also 2, but the character before those two 2's is either a 0 or a 1. Then any admissible sequence of length $n - 4$ is at the start. So the recurrence is $s_n = 3s_{n-1} - 2s_{n-4}$. It follows that

$$s_5 = 3s_4 - 2s_1 = 3 \cdot 76 - 2 \cdot 3 = 228 - 6 = 222, \text{ and}$$

$$s_6 = 3s_5 - 2s_2 = 3 \cdot 222 - 2 \cdot 9 = 666 - 18 = 648.$$

- 5.** Use the algorithm developed in class to find an Euler circuit in the graph G shown below (use node 1 as root):



The algorithm we learned in class always takes the least integer which represents an admissible choice. So starting at 1, we would produce the following initial sequence:

$$(1, 7, 4, 3, 2, 6, 4, 9, 2, 11, 6, 10, 5, 7, 9, 3, 8, 1)$$

But at this point, there are no legal moves. Again, the algorithm we learned in class scans forward to find the first position from which a legal move can be made. This entry is 9. Now the algorithm would find the following sequence:

$$(9, 8, 14, 11, 9)$$

So this second sequence would be inserted into the first to form:

$$1, 7, 4, 3, 2, 6, 4, 9, 8, 14, 11, 9, 2, 11, 6, 10, 5, 7, 9, 3, 8, 1)$$

Now that we have exhausted all edges, the algorithm halts.

You should be aware that the algorithm can halt in three different ways:

1. An Euler circuit has been completed.
2. A subroutine ends at distinct vertices x and y . In this case x and y are two vertices of odd degree in G and the graph does not have an Euler circuit.
3. The subroutine ends, not all edges have been traversed but there are no vertices on the path through this stage which are incident with edges which are not yet traversed. In this case, the graph is disconnected and edges not yet traversed are in a different component.

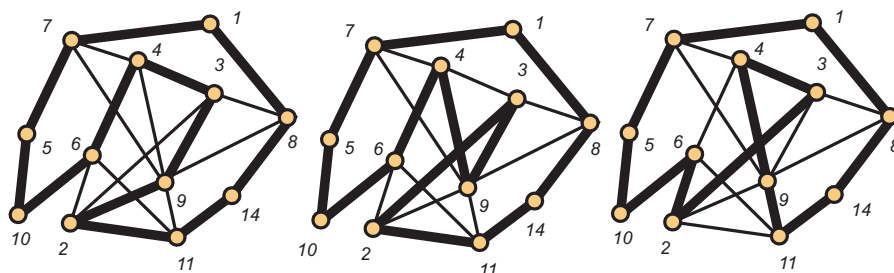
Programming Note. Implementing this algorithm requires us to keep track of the untraversed edges incident with each node. This raises interesting data structure issues. For example, at an intermediate point in the algorithm, we have to delete two edges incident with x , the entering edge and the departing edge. Exactly how this information is stored and how these updates are carried out is the programming challenge.

- 6.** Consider again the graph from the preceding problem. List the vertices in an order that shows why the graph is also hamiltonian.

A key ingredient to this problem is that we have no effective algorithm for finding a Hamiltonian cycle in a graph. On this test, we present a modest size graph where an exhaustive search can be carried out without much difficulty. For the given graph, there are (at least) three essentially different answers:

$$(1, 7, 5, 10, 6, 4, 3, 9, 2, 11, 14, 8) \quad (1, 7, 5, 10, 4, 9, 3, 2, 11, 14, 8) \quad (1, 7, 5, 10, 6, 2, 3, 4, 9, 11, 14, 8)$$

These are illustrated in the figure below.

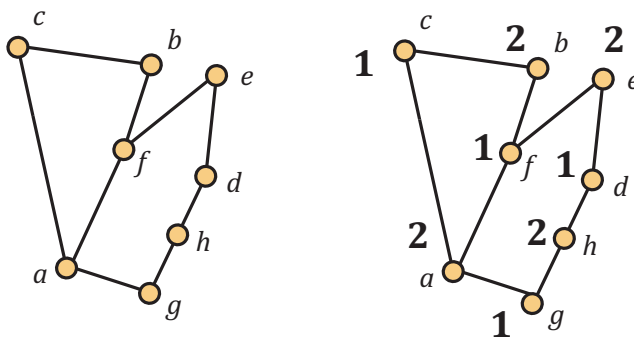


Of course, each of these cycles can be traversed in two directions. Also, any of the 14 vertices can be taken as the first point in the listing, but most people do not consider these as different cycles. Finally, note that while these three answers are certifiably correct, we have no effective way of determining whether or not there are more. In this case, it *appears* that these are the only correct answers.

Note that in listing the vertices that form a hamiltonian cycle, it is critical that the last vertex in the sequence be adjacent to the first, so that we are actually describing a *cycle*. There is also interest in the concept of a hamiltonian *path*, where this convention is dropped. For example, the following sequence is a Hamiltonian path but it is not a Hamiltonian cycle.

$$(1, 7, 5, 6, 2, 9, 11, 14, 8, 3, 4)$$

7. Consider the graph G shown on the left below.



a. Determine $\omega(G)$.

It is easy to see that $\omega(G) = 2$, since any adjacent pair of vertices forms a clique of size 2. However, there are no triangles (cliques of size 3).

b. Show that $\chi(G) = \omega(G)$ by providing a proper coloring of G . You may indicate your coloring by writing directly on the figure on the right.

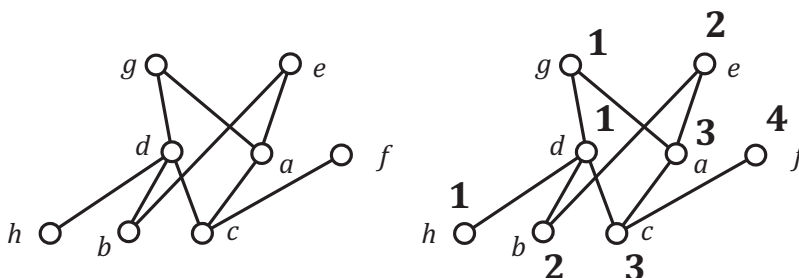
An appropriate coloring is provided. However, you should be aware that while detecting whether a graph is 2-colorable is *easy*, it is apparently quite difficult to answer whether it is t -colorable, when $t \geq 3$. So on any test or assignment, if you are asked to find $\chi(G)$ when the answer is three or more, you will not be following any algorithm known today. Skilled programmers have developed heuristics which seem to give reasonable results, but these are not exact algorithms and no certificates can be provided for their correctness. Remember, that for example, there are checkable certificates for a “yes” answer to the question: Is $\chi(G) \leq 3$? However, nobody knows how to provide a certificate for the correctness of a “no” answer to this question. Also, remember that if you discover such an algorithm, be sure to tell me about it first, and we can become famous together. And being terribly generous, I’ll be happy to share with you the \$1,000,000 cash award that comes with settling $P = NP$?. Who says that math doesn’t pay?

c. Explain why the graph G is perfect.

A graph is perfect if and only if $\chi(H) = \omega(H)$ for every induced subgraph H of G . In this case, when H is an induced subgraph and H contains an adjacent pair of vertices, then $\chi(H) = \omega(H) = 2$. On the other hand, if H does not contain an adjacent pair of vertices, then $\chi(H) = \omega(H) = 1$.

More generally, what we have actually shown is that any graph G with $\omega(G) \leq 2$ is perfect. Quite recently, Chudnovsky (Columbia), Robertson (Ohio State), Seymour (Princeton) and Thomas (Georgia Tech) completed the proof of the following deep theorem: A graph G is perfect if and only if neither G nor the complement of G contains an odd cycle of 5 or more vertices as an induced subgraph. The result was conjectured more than 40 years ago by the late Claude Berge, and the proof (more than 200 pages) appears in the *Annals*, arguably one of the world's most prestigious mathematics journals.

8. Consider the poset P shown below on the left.



a. Which of the ordered pairs in the following list belong to the reflexive, antisymmetric and transitive binary relation which defines this poset.

$$(d, b) \quad (h, e) \quad (a, e) \quad (f, f) \quad (c, g) \quad (e, b) \quad (b, b)$$

The binary relation defining the poset consists of all pairs $(x, y) \in X \times X$ with $x \leq y$ in P . So the answer becomes:

$$(a, e) \quad (f, f) \quad (c, g) \quad (b, b)$$

b. The poset P is not an interval order. By inspection, find four points which determine a subposet isomorphic to $\mathbf{2} + \mathbf{2}$.

The answer is $\{a, d, e, h\}$. Note that in a subsequent problem, we study an algorithm for finding a representation for an interval order—using the least number of end points. This algorithm will detect the presence of a copy of $\mathbf{2} + \mathbf{2}$, but one can do this in $O(n^4)$ time, just from the basic definition.

c. What is the width of the poset P ?

The width of this poset is 4. The teaching point here is that at this stage of our course, we had learned Dilworth's theorem: A poset of width w can be partitioned into w chains. However, given a finite poset P , we did not have an effective algorithm for finding the width. And of course, we did not have a method for finding a minimum partition of the poset into chains. These issues were deferred until the material on network flows and the follow-up application to bipartite matchings had been presented

d. List a set of elements which forms a maximum antichain in P .

$\{a, b, f, h\}$ is a 4-element antichain in P . Again this was done by inspection, not by following a specified algorithm.

e. Find a Dilworth partition of the poset P . You may provide your answer by writing directly on the figure.

Returning to the figure above, the drawing on the right accomplishes this task. Once more, we stress that this partition was found by inspection.

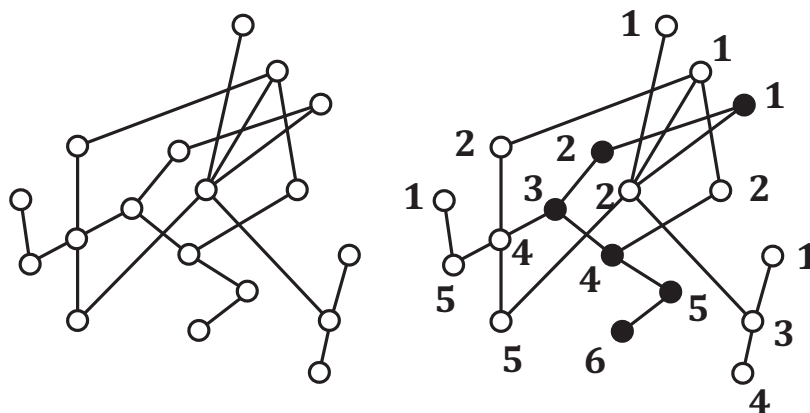
9. For the subset lattice 2^{17} , the statements listed below have been completed to form answers to questions on the test.

- a. The total number of elements is: 2^{17} . Maple says this is 131072.
- b. The total number of maximal chains is: $17!$. Maple says this is 355687428096000.
- c. The number of maximal chains through $\{3, 5, 8, 11\}$ is: $4!11!$. Maple says this is 958003200.
- d. The width of 2^{17} is: $\binom{17}{8}$. Of course, this answer can also be written as $\binom{17}{9}$. Maple reports this is 24310.

The preceding questions are central to the proof of Sperner's theorem: The width of the subset lattice 2^n is $\binom{n}{\lfloor n/2 \rfloor}$. The subset lattice is *ranked*, i.e., all maximal chains are maximum, and the width is just the maximum rank size.

As a side note, we showed that when $n \geq 2$, the cover graph of the poset 2^n is hamiltonian.

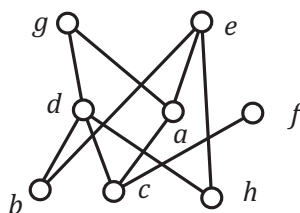
10. For the poset P shown on the left below, find the height h and a partition into h antichains by recursively stripping off the set of maximal elements. You may display your answer by writing directly on the diagram. Then darken a set of points that form a maximum chain.



The height of this poset is 6. Again, returning to the figure above, the drawing on the right illustrates the prescribed partition, while the darkened points form a maximum chain.

It should be noted that in general, when the height of P is h , there may be many different partitions into h antichains. For example, one could just as well recursively strip off the sets of minimal elements (on the final exam, several students did just that). However, we know that there cannot be a partition into fewer than h antichains.

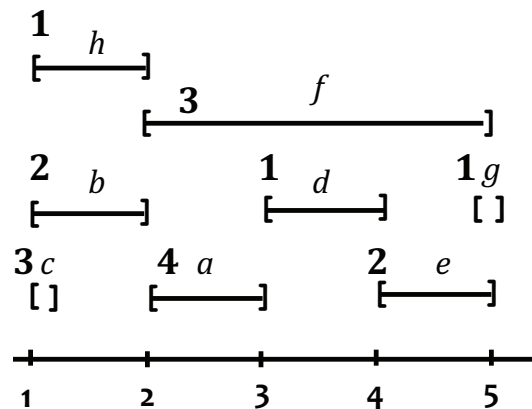
11. The poset P shown below is an interval order:



a. Find the down sets and the up sets. Then use these answers to find an interval representation of P that uses the least number of end points. *Note.* The answers for the first two parts have already been entered. Of course, on the exam, these spaces were blank.

$D(a) = \{c\}$	2	$U(a) = \{e, g\}$	3	$I(a) = [2, 3]$
$D(b) = \emptyset$	1	$U(b) = \{d, e, g\}$	2	$I(b) = [1, 2]$
$D(c) = \emptyset$	1	$U(c) = \{a, d, e, f, g\}$	1	$I(c) = [1, 1]$
$D(d) = \{b, c, h\}$	3	$U(d) = \{g\}$	4	$I(d) = [3, 4]$
$D(e) = \{a, b, c, h\}$	4	$U(e) = \emptyset$	5	$I(e) = [4, 5]$
$D(f) = \{c\}$	2	$U(f) = \emptyset$	5	$I(f) = [2, 5]$
$D(g) = \{a, b, c, d, h\}$	5	$U(g) = \emptyset$	5	$I(g) = [5, 5]$
$D(h) = \emptyset$	1	$U(h) = \{d, e, g\}$	2	$I(h) = [1, 2]$

b. In the space below, draw the representation you have found. Then use the First Fit Coloring Algorithm for interval graphs to solve the Dilworth Problem for this poset, i.e., find the width w and a partition of P into w chains. You may display your answers by writing the colors directly on the intervals in the diagram.



c. Find a maximum antichain in P :

The 4-element set $\{a, b, f, h\}$ is a maximum antichain in P .

12 a. Write all the partitions of the integer 7 into odd parts:

There are five partitions of the integer 7 into odd parts:

$$\begin{aligned}
 7 &= 7 \\
 &= 5 + 1 + 1 \\
 &= 3 + 3 + 1 \\
 &= 3 + 1 + 1 + 1 + 1 \\
 &= 1 + 1 + 1 + 1 + 1 + 1 + 1
 \end{aligned}$$

b. Write all the partitions of the integer 7 into distinct parts:

There are five partitions of the integer 7 into distinct parts:

$$\begin{aligned}7 &= 7 \\ &= 6 + 1 \\ &= 5 + 2 \\ &= 4 + 3 \\ &= 4 + 2 + 1\end{aligned}$$

Of course, for any integer n , the number of partitions of n into odd parts is always equal to the number of partitions of n into distinct parts.

Note. It is a challenging problem to compute $\pi(n)$, the number of partitions of the integer n . You may wonder whether this is easier or harder than computing the then number $p(n)$ of partial orders on an n -element set, a problem discussed earlier. In fact, it is much easier. For example, here is the value of $\pi(1000)$.

$$\pi(1000) = 24061467864032622473692149727991$$

You might be curious to know how many partitions of 1,000 have distinct parts. Here is the answer:

$$\pi_d(1000) = 8635565795744155161506.$$

Of course, this is the same as the number of partitions of 1,000 into odd parts, but you will take some comfort in the fact that I won't ask to list them and verify this fact—that's what proofs are for!

Finally, I comment that with today's computing limits, it is inconceivable to me that anyone could compute $p(1000)$, the number of partial orders on an 1,000-element set.

13a. Find the general solution to the advancement operator equation:

$$(A - 2)^3(A + 5)^2(A - 3)(A + 7)f = 0$$

Standard problem. When $(A - r)^m$ is a factor, so that r is a root of multiplicity m , then we have the following basis vectors in the solution space:

$$r^n \quad nr^n \quad n^2r^n \quad \dots \quad n^{m-1}r^n$$

Note that when $r = 1$, so $(A - 1)^m$ is a factor, these terms are all polynomials.

So in this problem as posed, the general solution is:

$$f(n) = c_1 2^n + c_2 n 2^n + c_3 n^2 2^n + c_4 (-5)^n + c_5 n (-5)^n + c_6 3^n + c_7 (-7)^n.$$

b. Write the form of a particular solution of the non-homogeneous advancement operator equation (do not carry out the work necessary to evaluate any constants in your answer):

$$(A - 2)^3(A + 5)^2(A - 3)(A + 7)f = 4(3)^n$$

Since 3 is a root and $c3^n$ is part of the solution in the homogeneous case, we must instead try for $cn3^n$. Another way to see this is to observe that if we apply $(A - 3)$ to both sides, then the answer we seek is part of the solution to the new homogeneous problem, which now has a factor of $(A - 3)^2$.

c. Find the solution to the advancement operator equation:

$$(A^2 - 5A + 6)f(n) = 0, \quad f(0) = 3 \text{ and } f(1) = 14.$$

The polynomial $A^2 - 5A + 6$ factors as $(A - 2)(A - 3)$, which has roots 2 and 3. So the solution we seek has the form $f(n) = c_1 2^n + c_2 3^n$. Substituting $n = 0$, we see that $c_1 + c_2 = 3$. Substituting $n = 1$, we see that $2c_1 + 3c_2 = 14$. Solving these two equations results in $c_1 = -5$ and $c_2 = 8$, so our final answer is:

$$f(n) = -5 \cdot 2^n + 8 \cdot 3^n.$$

14a. Write the inclusion/exclusion formula for the number $S(n, m)$ of onto functions from $\{1, 2, \dots, n\}$ to $\{1, 2, \dots, m\}$.

The derivation of the formula for $S(n, m)$ was done in class. It is also detailed in the book. Recall that you were asked explicitly to commit this formula to memory:

$$S(n, m) = \sum_{k=0}^m (-1)^k \binom{m}{k} (m - k)^n.$$

b. Write the inclusion/exclusion formula for the number d_n of derangements on $\{1, 2, \dots, n\}$.

Again, this is something you are supposed to know:

$$d_n = \sum_{k=0}^n (-1)^k \binom{n}{k} (n - k)!.$$

c. Evaluate your formula for d_n when $n = 6$.

$$\begin{aligned} d_6 &= \sum_{k=0}^6 (-1)^k \binom{6}{k} (6 - k)! \\ &= \binom{6}{0} 6! - \binom{6}{1} 5! + \binom{6}{2} 4! - \binom{6}{3} 3! + \binom{6}{4} 2! - \binom{6}{5} 1! + \binom{6}{6} 0! \\ &= 1 \cdot 720 - 6 \cdot 120 + 15 \cdot 24 - 20 \cdot 6 + 15 \cdot 2 - 6 \cdot 1 + 1 \cdot 1 \\ &= 265. \end{aligned}$$

The number d_n of derangements can be readily computed for relatively large values of n , and you can find simple formulas for Maple and Mathematica on the web. Here is one specific value I found:

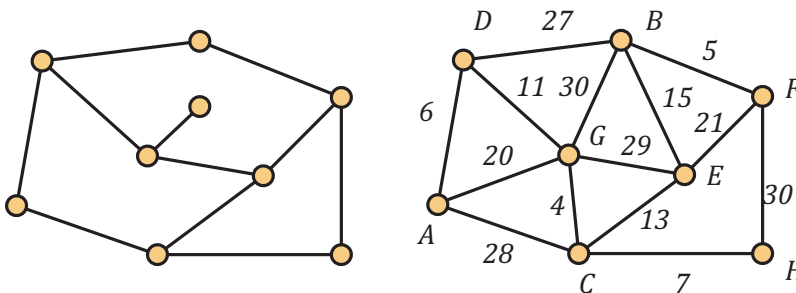
$$d_{20} = 895014631192902121.$$

d. Find the value of the Euler ϕ -function $\phi(n)$ when $n = 2^3 \cdot 5 \cdot 7^2$.

$$\begin{aligned} \phi(2^3 \cdot 5 \cdot 7^2) &= 2^3 \cdot 5 \cdot 7^2 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{5}\right) \left(1 - \frac{1}{7}\right) \\ &= (2^3 \cdot 5 \cdot 7^2) \frac{1}{2} \frac{4}{5} \frac{6}{7} \\ &= 2^2 \cdot 7 \cdot 4 \cdot 6 \\ &= 672. \end{aligned}$$

One of my favorite examples is to talk about finding $\phi(n)$ when it is known that $n = p \cdot q$, where p and q are large primes. When the number of digits is several hundred and you know p and q explicitly, you can find $\phi(n)$ very, very quickly. But if you don't know these values and you are just given n , then you are in a world of hurt. You have little chance of factoring n and no chance of using any kind of loop to determine $\phi(n)$ from the original definition.

15a. Verify Euler's formula for the planar graph shown on the left below.



Euler's formula asserts that when a connected planar graph G is drawn without edge crossings in the plane, then $V - E + F = 2$, where V is the number of vertices, E is the number of edges and F is the number of faces. In this case, $V = 9$, $E = 11$ and $F = 4$ (don't forget to count the exterior face), so $9 - 11 + 4 = 2$, as required.

b. Now consider the weighted graph shown on the right above. In the space below, list *in order* the edges which make up a minimum weight spanning tree using, respectively Kruskal's Algorithm (avoid cycles) and Prim's Algorithm (build tree). For Prim, use vertex A as the root.

Kruskal's Algorithm

- CG 4
- BF 5
- AD 6
- CH 7
- DG 11
- CE 13
- BE 15

Prim's Algorithm

- AD 6
- DG 11
- CG 4
- CH 7
- CE 13
- BE 15
- BF 5

This particular problem does not really illustrate what is going on with Kruskal. If you sort the edges by weight, as many students did on the final exam, then the spanning tree is just the first seven edges in the sorted list. In general, Kruskal always takes the two cheapest edges, but after that, it may be necessary to skip an edge when it forms a cycle when added to the edges already selected.

On the other hand, this example does serve to illustrate Prim fairly well. Note that the two cheapest edges are taken well into the process.

Programming types should be sensitive to the implementation details of these two algorithms. Kruskal seems to require a preliminary sort of the edges so if the graph has n vertices and q edges, the first step (sorting) takes $q \log q$ steps. Furthermore, as the algorithm proceeds, you need to keep track of whether the two end points of an edge are in different components of the forest being assembled.

In implementing Prim, you need to be able to quickly identify the cheapest edge with exactly one end point in the partial tree built to this stage.

16. Consider again the weighted graph from the preceding problem. Consider the weights as lengths, with all edges capable of being traversed in either direction. Apply Dijkstra's algorithm to find the distance from vertex A to all other vertices in the graph. Also, for each vertex X , find a shortest path from A to X .

We start with the trivial paths:

$$\begin{aligned} P(A, B) &= (A, B) & d &= \infty \\ P(A, C) &= (A, C) & d &= 28 \\ P(A, D) &= (A, D) & d &= 6 \\ P(A, E) &= (A, E) & d &= \infty \\ P(A, F) &= (A, F) & d &= \infty \\ P(A, G) &= (A, G) & d &= 20 \\ P(A, H) &= (A, H) & d &= \infty \end{aligned}$$

Of these distances, the smallest is 6 so the path $P(A, D)$ is declared permanent. Next we scan from D to see if any paths can be improved. This results in improvements for B and G :

$$\begin{aligned} P(A, B) &= (A, D, B) & d &= 6 + 27 = 33 < \infty \\ P(A, C) &= (A, C) & d &= 28 \\ P(A, E) &= (A, E) & d &= \infty \\ P(A, F) &= (A, F) & d &= \infty \\ P(A, G) &= (A, D, G) & d &= 6 + 11 = 17 < 20 \\ P(A, H) &= (A, H) & d &= \infty \end{aligned}$$

Now the shortest of these paths is $P(A, G) = (A, D, G)$ so this declared permanent. Next we scan from G , looking for improvements. We get them for C and E .

$$\begin{aligned} P(A, B) &= (A, D, B) & d &= 33 \\ P(A, C) &= (A, D, G, C) & d &= 17 + 4 = 21 < 28 \\ P(A, E) &= (A, D, G, E) & d &= 17 + 29 = 46 < \infty \\ P(A, F) &= (A, F) & d &= \infty \\ P(A, H) &= (A, H) & d &= \infty \end{aligned}$$

Now the shortest path is $P(A, C) = (A, D, G, C)$ so this is marked permanent. We then scan from C and get improvements for E and H .

$$\begin{aligned}
P(A, B) &= (A, D, B) \quad d = 33 \\
P(A, E) &= (A, D, G, C, E) \quad d = 13 + 21 = 34 < 46 \\
P(A, F) &= (A, F) \quad d = \infty \\
P(A, H) &= (A, D, G, C, H) \quad d = 21 + 7 = 28 < \infty
\end{aligned}$$

The shortest path is $P(A, H)$ so this becomes permanent and we scan from H . We get an improvement for F .

$$\begin{aligned}
P(A, B) &= (A, D, B) \quad d = 33 \\
P(A, E) &= (A, D, G, C, E) \quad d = 34 \\
P(A, F) &= (A, D, C, C, H, F) \quad d = 30 + 28 = 58 < \infty
\end{aligned}$$

We mark $P(A, B)$ as permanent we scan from B . We get an improvement for F .

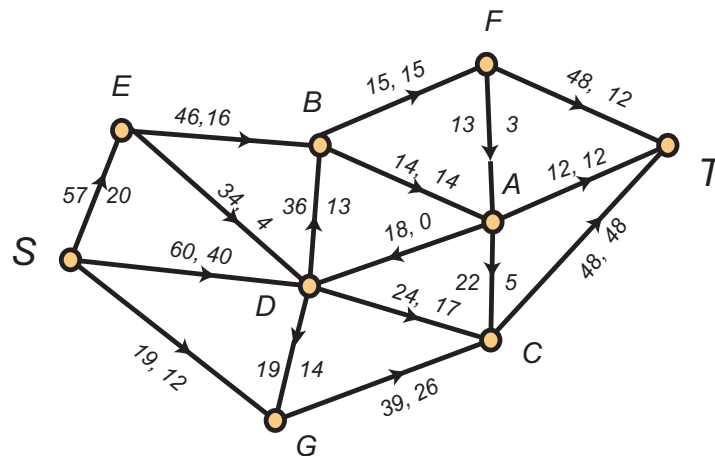
$$\begin{aligned}
P(A, E) &= (A, D, G, C, E) \quad d = 34 \\
P(A, F) &= (A, D, B, F) \quad d = 5 + 33 = 38 < 58
\end{aligned}$$

$P(A, E)$ is marked permanent and we scan from E but do not get an improvement for F . So finally, $P(A, F) = (A, D, B, F)$, with a total length of 38 is marked permanent.

There are many ways to organize this work. Here it has been done horizontally. In lectures I tend to do it horizontally. However, you CS types should recognize that this is a very space effective algorithm, since in an implementation, old paths may be discarded as new and improved versions are discovered. So the space requirements are minimal.

However, some students in the class just eye-balled the given graph and found shortest paths by inspection—not following Dijkstra at all. Not good. Lost points. Didn't show that you have an understanding of how this would be done for a graph on 1000 vertices with 400,000 edges.

17. Consider the following network flow:



a. What is the current value of the flow?

The current value of the flow is $20 + 40 + 12 = 72$, the amount of flow on the edges leaving the source. Note that $72 = 12 + 12 + 48$, the amount of flow arriving at the sink.

b. What is the capacity of the cut $V = \{S, B, D, E, G\} \cup \{A, C, F, T\}$.

To compute the capacity of the cut, we add up the capacity of all edges *from* to the first set *to* the second set. This results in

$$15 + 14 + 24 + 39 = 92.$$

Note that in making this calculation, we do not include the capacity of the edge from A to D .

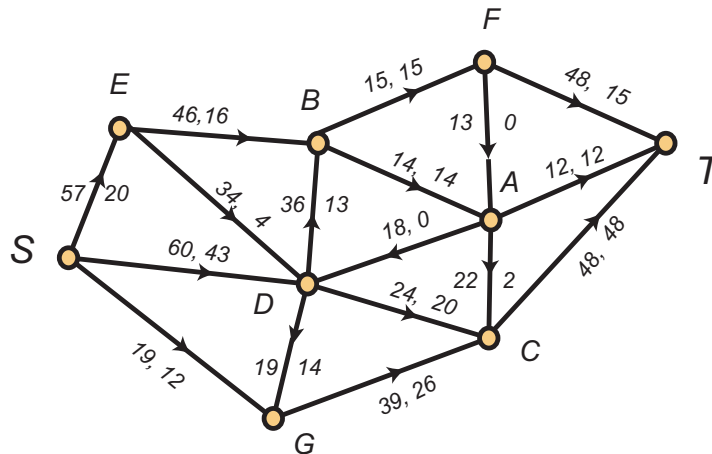
c. Carry out the labeling algorithm, using the pseudo-alphabetic order on the vertices and list below the labels which will be given to the vertices.

In order, the following labels are applied:

$$S(*, +, \infty), D(S, +, 20), E(S, +, 37), G(S, +, 7), B(D, +, 20), C(D, +, 7), A(C, -, 5), F(A, -, 3), T(F, +, 3).$$

d. Use your work in part c to find an augmenting path.

Backtracking we discover the augmenting path (S, D, C, A, F, T) on this path the edges (S, D) , (D, C) and (F, T) are forward while (C, A) and (A, F) are backwards. The flow is increased by 3 on the forward edges and decreased by 3 on the backwards edges. Of course, the value of the updated flow is $75 = 3 + 72$. The new flow is shown in the updated figure below.



e. Carry out the labeling algorithm a second time on the updated flow. It should halt without the sink being labeled.

In order, the following labels are applied:

$$S(*, +, \infty), D(S, +, 17), E(S, +, 37), G(S, +, 7), B(D, +, 17), C(D, +, 4), A(C, -, 2)$$

f. Find a cut whose capacity is equal to the value of the updated flow.

When the labelling algorithm halts, the labelled vertices are $\mathcal{L} = \{S, A, B, C, D, E, G\}$ and the unlabelled vertices are $\mathcal{U} = \{T, F\}$, so this is the desired cut. Note that all edges from

\mathcal{L} to \mathcal{T} are full, while all edges (there is only one) from \mathcal{U} to \mathcal{L} are empty. Accordingly, the capacity of this cut is $15 + 12 + 48 = 75$, which is the value of the current (updated) flow.

18. Consider a poset P whose ground set is $X = \{a, b, c, d, e, f, g, h, i\}$. Network flows (and the special case of bipartite matchings) are used to find the width w of P and a minimum chain partition. When the labelling algorithm halts, the following edges are matched:

$$e'h'' \quad c'f'' \quad d'e'' \quad h'b''$$

a. Find the chain partition of P that is associated with this matching.

Recall that when $x'y''$ is an edge in the matching then x is immediately under y in one of the chains in the matching. Also, recall that if a point x has the property that there are no matching edges of the form $x'y''$ and also, there are no matching edges of the form $z'x''$, then x is a one point chain in the partition. It follows that there are five chains in the partition:

$$C_1 = \{d < e < h < b\}$$

$$C_2 = \{c < f\}$$

$$C_3 = \{a\}$$

$$C_4 = \{g\}$$

$$C_5 = \{i\}$$

b. Find the width w of the poset P .

The width of P is the number of chains in the partition, which is five.

c. Explain why elements a , i and g belong to every maximum antichain in P .

The width of P is five, so every maximum antichain must intersect each of the five chains in the partition. Since C_3 , C_4 and C_5 are all one-element chains, each maximum antichain in P must contain a , g and i .

Although this detail is not reflected on the final exam, recall how the halting condition is used to find a maximum antichain. For each chain C in the final partition, there is a point x in C where x' is labelled and x'' is unlabelled when the Ford-Fulkerson algorithm halts. The set of all points determined in this manner always forms a maximum antichain.

I have always considered this result as a *capstone* event. First, we proved Dilworth's theorem in the classical style, much as was the custom forty years ago when nobody thought people would actually compute the width of a poset and find minimum chain partitions.

We studied network flows as a special class of linear programming problems and learned the essential property that when posed with integral capacities, there is always an integer valued maximum flow.

This integrality property is particularly useful when edges have capacity 1, since the Ford-Fulkerson algorithm will then only use the values 0 and 1 for flows on edges. A value of 1 is interpreted that the edge is "taken" and a value of 0 means that an edge is "not taken". This perspective allows for a *combinatorial* interpretation of flows and applications, for example, to bipartite matchings. We mentioned briefly applications to graph theory, such as Menger's theorems, and those of you who take additional graph theory course work will learn about these results in greater detail.

19. True–False. The following True-False questions are answered (sometimes with comments) at the end. On the test, you were asked to mark them in the left margin.

1. $2^{40} > 100,000,000$. True. $2^{40} = 2^{4 \cdot 10}$ and $2^4 = 16 > 10$, so $2^{40} > 10^{10} > 100,000,000$.
2. There is a planar graph G on 328 vertices with $\chi(G) = 9$. False. The Four Color Theorem asserts that every planar graph can be colored with four colors, i.e., $\chi(G) \leq 4$ when G is planar.
3. All graphs with 986 vertices and 4073 edges are non-planar. True. The maximum number of edges in a planar graph on n vertices is $3n - 6$ when $n \geq 3$. This is an easy consequence of Euler's formula.
4. There is a non-hamiltonian graph on 684 vertices in which every vertex has degree 426. False. Dirac's theorem asserts that if G is a graph on n vertices and all vertices have degree at least $\lceil n/2 \rceil$, then G is hamiltonian. Note that while we do know of an efficient algorithm for testing whether a graph is hamiltonian, there are a number of elegant results providing sufficient conditions for a graph to be hamiltonian. Dirac's theorem is just such a result. It has many generalizations and extensions.
5. Every connected graph on 783 vertices in which every vertex has degree 12 has an Euler circuit. True. A connected graph has an Euler circuit if and only if all vertices have even degree.
6. A cycle on 548 vertices is a homeomorph of the complete bipartite graph $K_{2,2}$. True. Note that $K_{2,2}$ is just C_4 , a cycle on four vertices.
7. When $n \geq 3$, the shift graph S_n has $\binom{n}{2}$ vertices, and $\binom{n}{3}$ edges. Furthermore, $\chi(S_n) = \lceil \lg n \rceil$. True. Shift graphs are an important instance of triangle-free graphs with large chromatic number.
8. The number of lattice paths from $(0,0)$ to (n,n) which do not pass through a point above the diagonal is the Catalan number $\binom{2n}{n}/(n+1)$. True. Detailed explanation given in class lectures and in text.
9. Any modern computer can accept a file of 3,000 positive integers, each at most 5,000, and quickly determine whether 3,742 is the sum of two integers in the file. True. This is an $O(n)$ problem, when n is the number of integers in the file and the size of the numbers is bounded.
10. Any modern computer can accept a file of 3,000 positive integers, each at most 5,000, and quickly determine whether 385,742 is the product of two integers in the file. True. An $O(n^2)$ problem.
11. Any modern computer can accept a file of 3,000 positive integers, each at most 5,000, and quickly factor each of the numbers into primes. True. This is an $O(n)$ problem—although the constant depends on the amount of time required to factor integers in the file. A modern computer can factor numbers of size at most 5,000 in less than a heartbeat. They run into problems with numbers having more than 100 digits, but the numbers we're talking about here have only four digits. Piece of cake.

12. There is a graph on 782 vertices in which no two vertices have the same degree. False. Argue by contradiction. Suppose G is such a graph. Then every vertex has degree at most 781. Since the degrees are all distinct, this means that for every i in $\{0, 1, 2, \dots, 781\}$, there is a unique vertex x_i in the graph with the degree of x_i being i . Now consider the vertex x_0 which then has degree zero, i.e., it is not adjacent to any other vertex in G . Then no vertex in G can have degree 781.
13. There is a poset with 723 points having width 69 and height 9. False. Consider a partition of P into 9 antichains. Since P has 723 points, the pigeon-hole principle requires one of the antichains to have at least $723/9$ points. But $723/9 > 80 > 69$, which is the width of P .
14. There is a sequence of 923 distinct positive integers which does not have an increasing subsequence of size 21 nor a decreasing subsequence of size 41. False. The Erdős-Szekeres theorem asserts that any sequence of $mn + 1$ distinct numbers has either an increasing subsequence of size $m + 1$ or a decreasing subsequence of size $n + 1$. Apply this result with $m = 40$ and $n = 20$, noting that $mn = 800$.
15. The permutation $(7, 1, 3, 5, 2, 4, 6)$ is a derangement. False. The third entry is a 3.
16. The number of equivalence relations on a set of size 100 is less than 10,000. False. Consider only those equivalence relations where every class has size 2 with one element from $\{1, 2, \dots, 50\}$ and the other from $\{51, 52, \dots, 100\}$. Evidently, there are $50!$ such equivalence relations, and $50! > 13 \cdot 12 \cdot 11 \cdot 10 > 10^4 = 10,000$.

Previously, we have commented on several counting problems. The problem of computing $\Pi(n)$, the number of equivalence relations on an n -element set (which is also called the number of partitions of an n -element set) has been computed for $n \leq 19$. Here are the known values:

$$\begin{aligned}
\Pi(1) &= 1 \\
\Pi(2) &= 3 \\
\Pi(3) &= 13 \\
\Pi(4) &= 73 \\
\Pi(5) &= 501 \\
\Pi(6) &= 4051 \\
\Pi(7) &= 37633 \\
\Pi(8) &= 394353 \\
\Pi(9) &= 4596553 \\
\Pi(10) &= 58941091 \\
\Pi(11) &= 824073141 \\
\Pi(12) &= 12470162233 \\
\Pi(13) &= 202976401213 \\
\Pi(14) &= 3535017524403 \\
\Pi(15) &= 65573803186921 \\
\Pi(16) &= 1290434218669921 \\
\Pi(17) &= 26846616451246353 \\
\Pi(18) &= 588633468315403843 \\
\Pi(19) &= 13564373693588558173
\end{aligned}$$

In this particular problem, we have been discussing $\Pi(100)$, a value which is far off the bottom of this table. But we did say that $\Pi(100) > 50!$. You might be interested to know just how large $50!$ actually is. Maple reports:

$$50! = 30414093201713378043612608166064768844377641568960512000000000000.$$

17. The binary relation $R = \{(a, a), (a, b), (b, b), (c, c)\}$ is reflexive on $X = \{a, b, c\}$. True. Just needs to have (x, x) for every $x \in X$.
18. The binary relation $R = \{(a, a), (a, b), (c, c), (b, c)\}$ is transitive on $X = \{a, b, c\}$. False. Has (a, b) and (b, c) . Needs to have (a, c) and it doesn't.
19. The binary relation $R = \{(a, a), (a, b), (c, c), (b, a)\}$ is antisymmetric on $X = \{a, b, c\}$. False. The relation contains (a, b) and (b, a) but $a \neq b$.
20. The binary relation $R = \{(a, a), (b, b), (c, c), (a, b), (b, a)\}$ is an equivalence relation on $X = \{a, b, c\}$. True. The relation is reflexive, symmetric and transitive.
21. Linear programming problems with integer coefficient constraints always have integer valued solutions. False. See the additional comments below.
22. Every linear programming problem is also a network flow problem. False. It's the other way round. A network flow problem is a special case of a linear programming problem.

As we remarked earlier, the fact that network flow problems with integer capacities has a maximum flow with all flow values integral is an *important* consequence of the Ford-Fulkerson algorithm.

Here is a trivial instance of a linear programming problem with integer constraints which does not have a solution in integers: minimize x subject to $3x = 5$. For a more illustrative example, consider: minimize $7x + 8y$ subject to the following constraints: (1) $9x + 4y \leq 36$; (2) $3x + 9y \leq 27$; (3) $x, y \geq 0$. The answer is the point in the plane where the two lines $9x + 4y = 36$ and $3x + 9y = 27$ cross. This point does not have integer coordinates (check it out yourself to be sure).