6A Graphs / Number theory

6A1 (Metric) A *distance function*, or a *metric*, in a set S is a function $d : (S \times S) \to \mathbb{R}$ such that for all $x, y, z \in S$:

(i) $d(x, y) \geq 0$ (nonnegativity)

(ii) $d(x, y) = 0$ if and only if $x = y$

(iii) $d(x, y) = d(y, x)$ (symmetry)

(iv) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)

Now we study distances between vertices of a graph.

(a) Prove or disprove: If $G = (V, E)$ is a connected graph, the following function ℓ is a metric in V:

$$\ell(x, y) = \text{the length of the shortest path between } x \text{ and } y$$

(b) Prove or disprove: If $G = (V, E)$ is a connected graph, the following function m is a metric in V:

$$m(x, y) = \ell(x, y) + 1$$

(Note: $m(x, y)$ is the number of vertices in a shortest path between x and y.)

(c) The *diameter* of a connected graph is the largest distance that can be found between any two vertices. What are the diameters of K_n (complete graph, $n \geq 1$), C_n (cycle, $n \geq 3$) and S_n (star graph, $n \geq 2$), when using ℓ as the distance function?

(d) List the diameters of C_n for $3 \leq n \leq 9$ and enter them to https://oeis.org/. How many different sequences are found? Look at the first search result. Do you think this is the diameter of C_n, looking at (i) its description and (ii) how the sequence continues?

Solution.

(a) Yes, ℓ is a metric in V.

Proof: Items (i),(ii),(iii) follow directly from the definition of ℓ. For (iv): Given a path p_1 from x to y of length $\ell(x, y)$ and a path p_2 from y to z of length $\ell(y, z)$, there exists a path from x to z of length $\ell(x, y) + \ell(y, z)$, by first going from x to y via p_1, then from y to z via p_2. Therefore $\ell(x, z) \leq \ell(x, y) + \ell(y, z)$.

(b) No. Item (ii) is not satisfied: If $x = y$ then $m(x, y) = \ell(x, y) + 1 = 0 + 1 = 1 \neq 0$.

(c) The diameter of K_n is 1, since there exists an edge between any two vertices by definition of K_n.

Th diameter of C_n is $\lfloor n/2 \rfloor$. That is, $n/2$ if n is even, otherwise $(n-1)/2$.

Proof: Clearly there exists vertices x, y in C_n such that $\ell(x,y) = \lfloor n/2 \rfloor$: From any x, the vertex y obtained by going through exactly $\lfloor n/2 \rfloor$ edges along the cycle. It remains to show that $\ell(x,y)$ cannot be greater than $\lfloor n/2 \rfloor$. For any x, y, given a path from x to y of length $p \leq n$, there is another of length $n - p$, by going in the other direction of the cycle. Suppose $\ell(x,y) > \lfloor n/2 \rfloor$, then there is another path of length $n - \ell(x,y) < n - \lfloor n/2 \rfloor = \lceil n/2 \rceil \leq \ell(x,y)$, contradicting that $\ell(x,y)$ is the length of the shortest path.

The diameter of S_n is 2.

Proof: Call the central vertex c. For any vertices $x, y \neq c$, there exists an edge from x to c and one from c to y, and no edge between x, y, by definition of S_n. Therefore $\ell(x,y) = 2$. Otherwise, for any two vertices x, y where at least one equals c, $\ell(x,y) \leq 1$.

(d) Write $D(n)$ the diameter of C_n. From (c), we have

$$D(3) = 1,\ D(4) = D(5) = 2,\ D(6) = D(7) = 3,\ D(8) = D(9) = 4.$$

Entering to https://oeis.org/, 427 results are found. The first result is the sequence of the number of primes $\leq n$, denoted by $\pi(n)$. This is not the diameter of C_n, since from (c) we know that $D(n) = \lfloor n/2 \rfloor$ for $n \geq 3$, i.e. each $n \in \mathbb{N}$ appears exactly twice in the sequence, which is not the case for $\pi(n)$.

6A2 (Disconnected graphs)

(a) Prove or disprove: In any graph, connected or disconnected, this function is a metric:

$$c(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \text{ and there is a path from } x \text{ to } y \\ 2 & \text{otherwise} \end{cases}$$

(b) Is ℓ (from problem 1) a well-defined metric in a disconnected graph?

Solution.

(a) Yes, c is a metric.

Proof: That Items (i),(ii),(iii) listed in Problem 6A1 are satisfied is direct from the definition of c. We show that (iv) is also satisfied.

Suppose there exists x, y, z such that (iv) does not hold. Since the range of c is $\{0, 1, 2\}$, there are only two possibilities.

(1) $d(x, z) = 1$ and $d(x, y) = d(y, z) = 0$. But the latter implies $x = y = z$, so that $d(x, z) = 0$, a contradiction. So this is not possible.

(2) $d(x, z) = 2$ and $d(x, y)$, $d(y, z)$ are either both 0, or one equals 1 and the other 0. That $d(x, y) = d(y, z) = 0$ is not possible for the same reason as above. Next, w.l.o.g. suppose $d(x, y) = 1$ and $d(y, z) = 0$. This implies there is a path from x to y, and $y = z$, so that there is a path from x to z, contradicting that $d(x, z) = 2$.

Since both cases yield contradiction, we conclude that for any x, y, z, the triangle inequality holds for c.

(b) No, because for x, y that are disconnected, the shortest path $\ell(x, y)$ between x and y is undefined.

6A3 (Family tree) Let $T_n = (V_n, E_n)$ be the (ancestral) family tree of a fixed person c, going back n generations ($n \geq 0$). T_0 has c as its only vertex. T_1 contains also the two parents of c, with edges from c to both. T_2 adds the four grandparents and so on.

(a) Find $|V_n|$, $|E_n|$, $\chi(T_n)$, and the diameter of T_n.

(b) Find the *average distance* from c to all people in the family tree,

$$a_n = \frac{1}{|V_n|} \sum_{x \in V_n} \ell(c, x),$$

when $n = 10$. Give the result in decimal form with three decimals.

Solution.

(a) $|V_n| = 1 + 2^1 + 2^2 + \ldots + 2^n = \frac{2^{n+1}-1}{2-1} = 2^{n+1} - 1$, where the second equality used the geometric series equality.

$|E_n| = 0$ if $n = 0$, otherwise $|E_n| = 2^1 + 2^2 + \ldots + 2^n = 2^{n+1} - 2$.

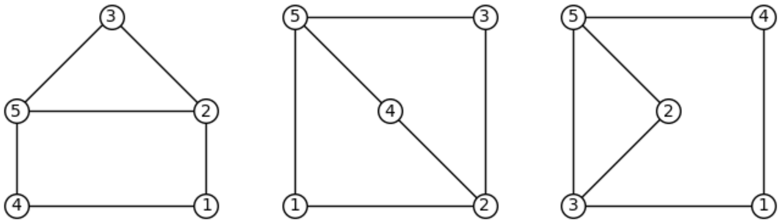
$\chi(T_n) = 2$, by alternating color in each layer of T_n (color 1 for vertices x such that $\ell(x, c)$ is even, otherwise color 2).

(b) In T_n, for any $i \in \{0, 1, \ldots, n\}$, the number of vertices $x \in V_n$ such that $\ell(c, x) = i$ is 2^i. Therefore

$$a_n = \frac{1}{2^{n+1} - 1} \sum_{i \in \{0, \ldots, n\}} (2^i \cdot i).$$

For $n = 10$, we have $a_{10} = 9.005$.

6A4 (Isomorphic or not) Consider the following three graphs. Find the degrees of each of their vertices. Find which graphs (if any) are isomorphic, and construct an isomorphism (function) between them. If two graphs are not isomorphic, explain why it is impossible to construct an isomorphism between them.



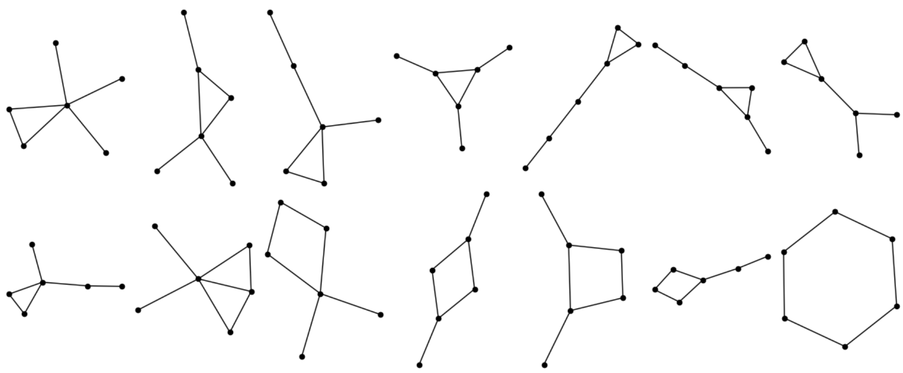
Solution. Call the left, middle and right graphs L, M, R respectively.

For L: The degrees of vertices $1, 2, 3, 4, 5$ are $2, 3, 2, 2, 3$ respectively.
For M: The degrees of vertices $1, 2, 3, 4, 5$ are $2, 3, 2, 2, 3$ respectively.
For R: The degrees of vertices $1, 2, 3, 4, 5$ are $2, 2, 3, 2, 3$ respectively.

L and R are isomorphic. Write V_L, V_R the vertex set of L, R respectively. Let $f : V_L \to V_R$, $f(3) = 2$, $f(2) = 3$, and $f(x) = x$ if $x \neq 2, 3$. Then f is an isomorphism from L to R. Bijectivity can be easily verified.

M is not isomorphic to L or R. Notice that in both L, R there exists a 3-cycle, but this is not the case in M.

6A5 (List correctness) A graph is *unicyclic* if it contains exactly one cycle graph (as a subgraph). Lecturer K. has constructed what he thinks is a complete listing of all unlabeled connected unicyclic graphs of six vertices.



As his research assistant you have learned to suspect his listings. You check https://oeis.org/A001429, and you find that the correct count of such graphs is 13, while K's list contains 14 graphs. Something is wrong. (Of course it *could* be the OEIS entry, but let's assume it is correct.)

(a) Find one graph that should not be on the list at all.

(b) After removing that graph, the count is now correct. Does this prove that the list is correct?

(c) You suspect that the list has one graph twice (a duplicate). Find it.

(d) You also suspect that one graph is missing from the list. What is it?

Solution.

(a) The second graph (counting from left to right) in the second row should not be on the list. Obviously it contains more than one cycle.

(b) No. For example, it could be that another graph in the list is also incorrect, while a correct graph is missing.

(c) The third graph in the first row is same as the first graph in the second row.

(d) The graph with a 5-cycle is missing.

6A6 (Divisibility) Recall that for two integers a, b (not necessarily positive!) we say that a *divides* b, or in short $a \mid b$, if there exists an integer m such that $b = ma$.

Prove or disprove each of the following statements. All variables are understood to be integers. For those statements that are true, *also* give a small example with concrete numbers (preferably positive). For the statements that are false, a small counterexample can probably be found.

(a) If $a \mid b$ and $b \mid c$, then $a \mid c$.

(b) If $a \mid b$, then $b \mid a$.

(c) If $a \mid b$ and $b \mid a$, then $a = b$ or $a = -b$.

(d) If $a \mid b$ and $a \mid c$, then $a \mid b + c$.[1]

(e) If $a \mid b$ and $a \mid b + c$, then $a \mid c$.

(f) If $a \mid b$, and c is any integer, then $a \mid bc$.

(g) If $a \mid c$ and $b \mid c$, then $ab \mid c$.

(h) If $a \mid c$ and $b \mid c$, then $ab \mid c^2$.

(i) If $a \mid c$ and $b \mid d$, then $ab \mid cd$.

(j) If $a = bc > 0$, then $b \le \sqrt{a}$ or $c \le \sqrt{a}$ (or both).

[1] You should read "$|$" as a relation symbol like "$=$" and "$<$", so it binds more loosely than any arithmetic operation. That is, $a \mid b + c$ means $a \mid (b + c)$.

Solution.

(a) True. Example: $2 \mid 4$ and $4 \mid 8$, then $2 \mid 8$.

Proof: If $a \mid b$ and $b \mid c$, there exists $k_1, k_2 \in \mathbb{Z}$ such that $k_1 a = b$ and $k_2 b = c$. Then $k_1 k_2 a = c$, so $a \mid c$.

(b) False. Counterexample: $3 \mid 6$, but $6 \nmid 3$.

(c) True. Example: $3 \mid -3$ and $-3 \mid 3$, then $3 = -(-3)$.

Proof: If $a \mid b$ and $b \mid a$, there exists $k_1, k_2 \in \mathbb{Z}$ such that $k_1 a = b$ and $k_2 b = a$. Then $a = k_1 k_2 a$, and $1 = k_1 k_2$. The only possibilities are $k_1 = k_2 = 1$ or $k_1 = k_2 = -1$, that is, $a = b$ or $a = -b$.

(d) True. Example: $2 \mid 4$ and $2 \mid 6$, then $2 \mid 10$.

Proof: If $a \mid b$ and $a \mid c$, there exists $k_1, k_2 \in \mathbb{Z}$ such that $k_1 a = b$ and $k_2 a = c$. Then $k_1 a + k_2 a = (k_1 + k_2)a = b + c$, so $a \mid b + c$.

(e) True. Example: $2 \mid 4$ and $2 \mid 12$, then $2 \mid 8$.

Proof: If $a \mid b$ and $a \mid (b + c)$, there exists $k_1, k_2 \in \mathbb{Z}$ such that $k_1 a = b$ and $k_2 a = b + c$. Then $k_2 a - k_1 a = (k_2 - k_1)a = c$, so $a \mid c$.

(f) True. Example: $2 \mid 4$, then $2 \mid 4c$ for any integer c.

Proof: If $a \mid b$, there exists $k \in \mathbb{Z}$ such that $ka = b$. Then $(kc)a = bc$, where $kc \in \mathbb{Z}$ because $c \in \mathbb{Z}$. So $a \mid bc$.

(g) False. Counterexample: $4 \mid 12$ and $6 \mid 12$, but $24 \nmid 12$.

(h) True. Example: $4 \mid 12$ and $6 \mid 12$, and $24 \mid 144$.

Proof: If $a \mid c$ and $b \mid c$, there exists $k_1, k_2 \in \mathbb{Z}$ such that $k_1 a = c$ and $k_2 b = c$. Then $(k_1 a)(k_2 b) = (k_1 k_2)ab = c^2$, so $ab \mid c^2$.

(i) True. Example: $2 \mid 4$ and $5 \mid 15$, then $10 \mid 60$.

Proof: If $a \mid c$ and $b \mid d$, there exists $k_1, k_2 \in \mathbb{Z}$ such that $k_1 a = c$ and $k_2 b = d$. Then $k_1 k_2 ab = cd$, so $ab \mid cd$.

(j) True. Example: $16 = 2 \cdot 8$, then $\sqrt{16} = 4 \geq 2$.

Proof: Suppose not, $b, c > \sqrt{a}$, then $bc > (\sqrt{a})^2 = a$, contradicting with $a = bc$. Therefore $b, c > \sqrt{a}$ cannot be true, in other words $b \leq \sqrt{a}$ or $c \leq \sqrt{a}$, or both.

6A7 (Simple Diophantus) Consider this equation in two variables.

$$5x - 2y = 1$$

By high-school methods you can easily find an infinite number of solutions $(x, y) \in \mathbb{R}^2$, indeed for any $x \in \mathbb{R}$ you could find a suitable $y \in \mathbb{R}$. However, we are now looking for integer solutions $(x, y) \in \mathbb{Z}^2$.

(a) Find one solution $(x, y) \in \mathbb{Z}^2$. (Hint: Trial and error with small positive integers should be enough.)

(b) From your one solution, construct another (x', y'). (Hint: If x increases by some amount c, how must y change so that the equality stays? Make it so that the new values are again integers.)

(c) Construct an infinite set of integer solutions (x, y) to the equation. (You do not need to prove that your construction contains *all* possible solutions, but at least it should contain infinitely many.)

Equations like this, where only integer solutions are desired, are called *Diophantine equations*, after Diophantus of Alexandria who lived in the 3rd century AD. His book *Arithmetica* contains many such problems.

Solution.

(a) An example solution is $(1, 2)$. It holds $5 \cdot 1 - 2 \cdot 2 = 1$.

(b) Suppose x is increased by 1, then y would need to increase by $5/2$ to maintain the equality, but $5/2 \notin \mathbb{Z}$. Then, one can instead increase x by 2 and y by $10/2 = 5$, where $2, 5 \in \mathbb{Z}$.

From $(1, 2)$, this yields another solution of (x, y) being $(3, 7)$.

(c) Continuing with the same reasoning as in (b), one can construct a new solution by adding(/subtracting) $(2, 5)$ to(/from) an existing one. This yields the set of solutions

$$\{(x, y) : x = 1 + 2k,\ y = 2 + 5k,\ k \in \mathbb{Z}\}.$$

Verifying, one has $5x - 2y = 5(1 + 2k) - 2(2 + 5k) = 1$, as desired. The above is an infinite set, because $\mathbb{Z}$ is.

6A8 (** CHALLENGE, worth an extra point: Variant of geometric series)

In discrete mathematics an often recurring task is computing sums.[2] One example is the *arithmetico-geometric sum*

$$s_n = \sum_{k=1}^{n} k a^k,$$

where $a \neq 1$ is a constant. Your task is to find a closed form expression[3] for it. Here are some hints to get started:

[2] Somewhat analogous to integrals in calculus — and sometimes quite as difficult.

[3] An expression containing a fixed, finite number of elementary arithmetic operations — in particular, not using the big sum symbol $\sum$ or three dots.

- Write out as_n as a sum.
- Write out $s_n - as_n$ as a sum, collecting like terms from s_n and as_n (terms having the same power of a).
- Solve for s_n, simplify, and recognize a part of your expression as something you know already.

Once you have your general closed form expression for s_n with any $a \neq 1$, write a specific form for s_n when $a = 2$, and try to simplify it further. Then verify it against explicit summation with $n = 1, 2, 3, 4$. Then calculate 3b again with your new shiny formula.

Calculate also s_n with $a = 10$ and $n = 5$ and verify against direct summation.

Finally, think why we had to assume $a \neq 1$ for our general method. If $a = 1$, do you know a simpler method?

Solution. We have

$$\begin{aligned} s_n &= a + 2a^2 + 3a^3 + \ldots (n-1)a^{n-1} + na^n \\ as_n &= a^2 + 2a^3 + 3a^4 + \ldots (n-1)a^n + na^{n+1} \end{aligned}$$

and subtracting the second equation from the first,

$$\begin{aligned} s_n - as_n &= a + (2-1)a^2 + (3-2)a^3 + \ldots + (n-(n-1))a^n - na^{n+1} \\ &= a + a^2 + a^3 + \ldots + a^n - na^{n+1} \\ &= \frac{a(1-a^n)}{1-a} - na^{n+1}, \end{aligned}$$

thus

$$s_n = \frac{a(1-a^n)}{(1-a)^2} - \frac{na^{n+1}}{1-a}.$$

With $a = 2$ this becomes

$$\begin{aligned} s_n &= \frac{2(1-2^n)}{(1-2)^2} - \frac{n2^{n+1}}{1-2} \\ &= 2(1-2^n) + n2^{n+1} \\ &= 2 + (n-1)2^{n+1}. \end{aligned}$$

With $a = 2$ and $n = 1, 2, 3, 4$ we get $s_n = 2, 10, 34, 98$.

For 3b we now have (using $a = 2$ and $n = 10$)

$$a_n = \frac{2 + 9 \cdot 2^{11}}{2^{11} - 1} = \frac{18434}{2047} \approx 9.005.$$

With $a = 10$ and $n = 5$ we get $s_n = 543210$, which matches the explicit summation

$$1 \cdot 10 + 2 \cdot 10^2 + 3 \cdot 10^3 + 4 \cdot 10^4 + 5 \cdot 10^5 = 10 + 200 + 3000 + 40000 + 500000.$$

Our general method contains division by $1 - a$ so it does not work with $a = 1$. But with $a = 1$ our sum is simply the sum of a finite arithmetic progression

$$\sum_{k=0}^{n} k = 1 + 2 + \ldots + k = \frac{k(k+1)}{2}.$$

(seen on lectures / exercise 4B4 (b).)

Our initial trick, multiplying the sum with a and subtracting from the original sum, may seem “out of the blue sky”. However, similar tricks are often used in simplifying difficult sums — usually the hope is that some things cancel, and we end up with a simpler or previously known sum.

This trick is (very loosely) analogous to “integration by parts” in calculus.

Well beyond the course requirements: For those who are not faint at heart, and want extra challenges, try if you can pull off the same trick for the sum

$$\sum_{k=0}^{n} k^2 a^k,$$

or with some higher power of k. No, we are not providing the answer here. (You can calculcate a few small sums directly and then try an OEIS lookup.)