Master of Data Science Online Programme
Course: Discrete Mathematics
SGA #2: Graphs

Student: Andrei Batyrov (Fall 2022)
arbatyrov@edu.hse.ru

Higher School of Economics (HSE)
National Research University
Faculty of Computer Science

October 17, 2022

Contents

List of Figures

Problem 1

In a country there are several airports. Airport A is directly connected to 23 other airports. Airport B has a direct connection to 3 other airports. Each airport, except A and B, is directly connected to 10 other airports. Prove that there is an airline route (maybe with flight changes) between A and B.

Solution

Following a similar problem solved in [2, pp. 25–26], let's try to prove the opposite: suppose there is no path between A and B. In this case, the country's airport network is a disconnected graph with two connected components for A and B. The first connected component is a graph $\mathcal{G}\langle V(\mathcal{G}) = \{ A, C_1, \ldots, C_{23}, \textit{other airports} \}, E(\mathcal{G}) = \{ e_1, \ldots, e_{23}, \textit{other routes} \}\rangle$. The second connected component is a graph $\mathcal{H}\langle V(\mathcal{H}) = \{ B, D_1, D_2, D_3, \textit{other airports} \}, E(\mathcal{H}) = \{ f_1, f_2, f_3, \textit{other routes} \}\rangle$. See Figure 1 for illustration. Edges going from/to other 10 airports for each child of A and B are not shown.

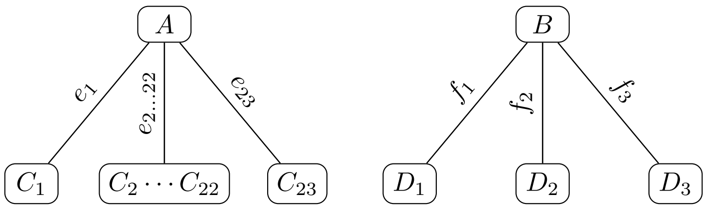


Figure 1: Airport network graph: two different connected components

As per [4, p. 27], degree of a vertex is the number of edges connected to it. Thus, $deg(A) = |E(\mathcal{G})| = 23$ and $deg(B) = |E(\mathcal{H})| = 3$. A vertex is called odd (even) if its degree is odd (even). So, both A and B are odd vertices.

As per the Handshaking Lemma [4, pp. 39–45], in a graph the number of vertices with odd degree (odd vertices) is even, or in other words, every graph has an even number of odd vertices. But each connected component will include exactly one odd vertex: A (and B), which is an odd number, which violates the Handshaking Lemma since we expect the number of odd vertices to be even. Thus, having two different connected components is a contradiction, i.e. vertices A and B must belong to the same connected component. In other words, the graph must be connected, so there must be a path from A to B. Since we have a condition that 'each airport, except A and B, is directly connected to 10 other airports', there is at least one path from A to B via their child vertices C_i and D_j.

Answer

There is an airline route (maybe with flight changes) between A and B.

Problem 2

Consider an Erdős–Rényi random graph on 4 vertices with $p = 1/2$. Calculate the probability that this graph is connected.

Solution

Let $\mathcal{G}(n, p)$ be an Erdős–Rényi random graph on n vertices with equal probability of drawing any edge independently p. As per [6], the probability of a concrete graph with n vertices and m edges to be in $\mathcal{G}(n, p)$ is

$$P(\mathcal{G}'(n,m) \in \mathcal{G}(n,p)) = p^m \cdot (1-p)^{\binom{n}{2}-m}. \tag{1}$$

Since $p = 1/2 = 1-p$, we can simplify (1) and confirm that $P(\mathcal{G}'(n,m) \in \mathcal{G}(n,p))$ does not depend on m: $P(\mathcal{G}'(n,m) \in \mathcal{G}(n,1/2)) = p^m \cdot p^{\binom{n}{2}-m} = p^{\binom{n}{2}}$. In our case $P(\mathcal{G}'(4,m) \in \mathcal{G}(4,1/2)) = (1/2)^{\binom{4}{2}} = (1/2)^6 = {}^1/_{64}$. So each concrete graph on 4 vertices is equally probable with $P(\mathcal{G}'(4,m) \in \mathcal{G}(4,1/2)) = \textit{ const } = {}^1/_{64}$.

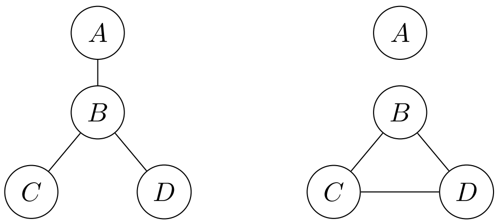
Let's consider a sample space Ω for our model, where each elementary outcome is one realization of a randomly drawn $\mathcal{G}(4, 1/2)$, i.e. a concrete graph $\mathcal{G}'(4,m)$. There are $2^{\binom{n}{2}} = 2^{\binom{4}{2}} = 64$ concrete graphs in our sample space, i.e. $|\Omega| = 64$. Let A be an event '$\mathcal{G}'(4,m)$ *is connected*'. Since the probability of each graph is equal, we can find the classical probability of event A as

$$P(A) = \frac{|A|}{|\Omega|}. \tag{2}$$

As per [2, p. 12], a graph is considered as connected, if any two vertices in it are connected, i.e. there is at least one path between any two vertices. In other words, all vertices of a connected graph must belong to the same one connected component. So, to find all outcomes in A, we need to find all connected graphs. To do so, we need to count the number of graphs for a given number of edges m and decide if such a graph can be in A. The simplest possible connected graph is a tree with $m = n - 1 = 4 - 1 = 3$ edges in it. However, a graph with 3 edges may have a cycle and thus be disconnected. To guarantee connectedness of a graph with unknown structure, the following condition must be met [2, p. 43]:

$$m \geq \frac{(n-1)\cdot(n-2)}{2}, \tag{3}$$

where n is the number of vertices and m is the number of edges. As per (3), $m \geq (4-1)\cdot(4-2)/2 = 3$. So, our graph must have at least three edges to be connected. If $m < 3$, then a graph is disconnected and is not in A. If $m > 3$, then the graph is guaranteed to be connected and is in A. Special case is with $m = 3$: a graph can be either connected (tree), or disconnected (has one cycle) – see Figure 2. So, $A = \{\, \mathcal{G}'(4,m) \mid 4 \leq m \leq 6 \,\} \cup \{\, \mathcal{G}'(4,3) \mid \textit{It is a tree} \,\}$.



$m_1 = 3$
tree
connected

$m_2 = 3$
one cycle
disconnected

Figure 2: $\mathcal{G}'(4, 3)$: connected and disconnected

The number of graphs $\mathcal{G}'(4,m)$ is the number of combinations of m from $\binom{n}{2} = \binom{4}{2} = 6$ maximal possible number of edges (complete graph), i.e. $\binom{6}{m}$. Let's count the number of graphs for each m we are interested in:

1. $m = 3$: $N_{\mathcal{G}'(4,3)} = \binom{6}{3} = 20$ (connected trees and disconnected graphs with one cycle);
2. $m = 4$: $N_{\mathcal{G}'(4,4)} = \binom{6}{4} = 15$ (connected graphs);

3. $m = 5$: $N_{\mathcal{G}'(4,5)} = \binom{6}{5} = 6$ (connected graphs);

4. $m = 6$: $N_{\mathcal{G}'(4,6)} = \binom{6}{6} = 1$ (connected complete graph).

For the special case with $m = 3$, from $N_{\mathcal{G}'(4,3)} = 20$ we need to exclude graphs with cycles and keep only trees. A cycle is a combination of 3 edges connecting any 3 vertices out of possible 4 edges – the fourth edge is not drawn and the the fourth vertex is not connected. It means that the number of such combinations is $\binom{4}{3} = 4$ graphs with one cycle. So, for $\mathcal{G}'(4, 3)$ we will have 20 (all such graphs) – 4 (only graphs with one cycle) = 16 trees. Now, $|A| = 15 + 6 + 1 + 16 = 38$. And finally, by (2): $P(\mathcal{G}(4, 1/2)\ \textit{is connected}) = P(\mathcal{G}'(4, m)\ \textit{is connected}) = 38/64 = 19/32 = 0.59375$.

Another Approach

A simulation of random experiment to draw concrete graphs $\mathcal{G}'(4, m)$ as realizations of $\mathcal{G}(4, 1/2)$ was done. Sequences of length of 6 ($m = 6$) were obtained with the NumPy Python library, where each sequence corresponds to outcomes of one Bernoulli process (coin toss simulation). Then each outcome was assigned to a corresponding edge (pair of vertices). A decision was made if a graph was connected or not based on m – the sum of all 6 binary outcomes. For example, for one Bernoulli process the outcomes were $\{0, 1, 1, 1, 0, 1\}$ which were mapped on edges as follows:

1. $AB = 0$ (edge not drawn)
2. $AC = 1$ (edge drawn)
3. $AD = 1$ (edge drawn)
4. $BC = 1$ (edge drawn)
5. $BD = 0$ (edge not drawn)
6. $CD = 1$ (edge drawn)

The sum of all edge mappings is 4, so the graph is considered as connected. For $m = 3$ cycles were excluded based on rules. The simulation was done for several number of batches each having several number of samples. For each batch $P(\mathcal{G}(4, 1/2)\ \textit{is connected})$ was estimated as the number of connected graphs $\mathcal{G}'(4, m)$ in the batch (over all samples in the batch) divided by the number of samples in the batch. The result of the simulation for 6 batches each with sample sizes of 100, 500, 1000, 5000, 10000, and 50000 is shown on Figure 3.

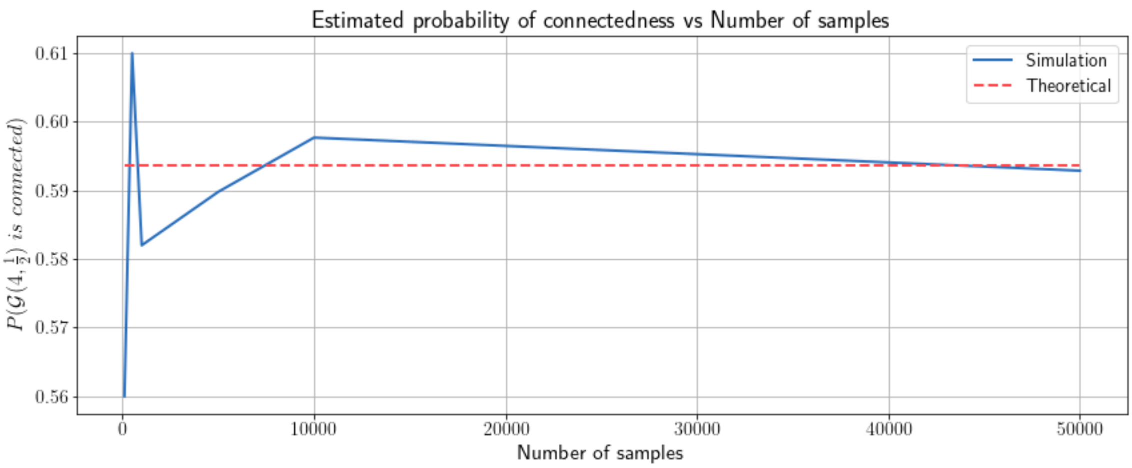


Figure 3: Estimation of $P(\mathcal{G}(4, 1/2)\ \textit{is connected})$ with computer simulation

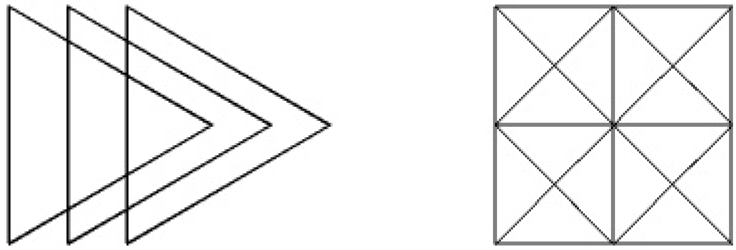
We can see that as the number of samples increases, the estimated $P(\mathcal{G}(4, 1/2)\ \textit{is connected})$ approaches the result obtained in Solution which is 0.59375.

Answer

The probability that this graph is connected is $19/32 = 0.59375$.

Problem 3

Which of the following pictures can be drawn in one stroke of pen, without traversing a line twice (like a Euler path in a graph)?



Solution

As per [3, pp. 16–19], an Euler path is a path in the graph which visits each edge exactly once and an Euler cycle is an Euler path which starts and ends at the same vertex. Euler's Theorem: a connected graph has an Euler cycle, if and only if every vertex has even degree.

We can easily see that the graph shown on left has no odd vertices since it only has vertices of degree 2 and 4. Also the graph is connected, so the Euler's Theorem holds, and we can walk the graph traversing each edge exactly once – see Figure 4. Both vertices and edges are enumerated. One of possible Euler paths and cycles with 21 edges: $1 \to 2 \to 3 \to 4 \to 5 \to 6 \to 7 \to 8 \to 9 \to 10 \to 11 \to 3 \to 12 \to 5 \to 13 \to 9 \to 15 \to 12 \to 14 \to 15 \to 11 \to 1$. Colors are used only for convenience to help walk the graph.

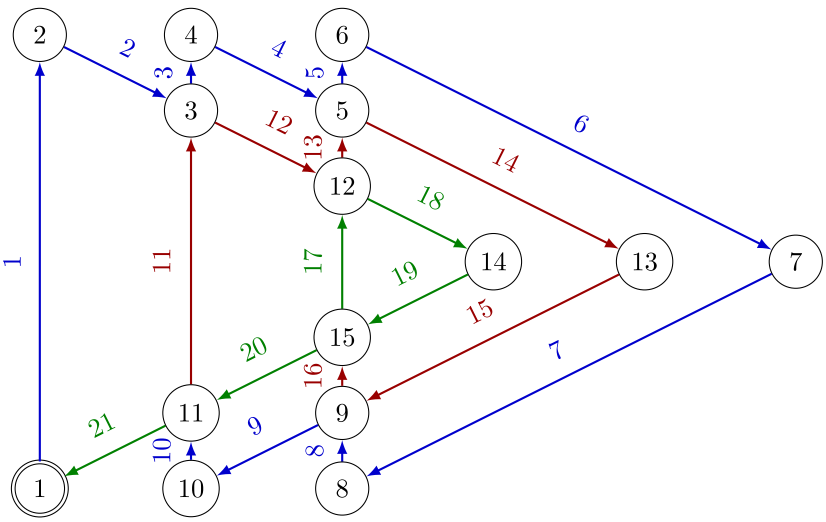


Figure 4: Euler path and cycle

Again, as per [3, p. 19], if there exists an Euler path, then the graph has at most two odd vertices. We can easily see that the graph shown on the right has 8 odd vertices: 4 vertices with degree of 3 (corners) and 4 vertices with degree of 5 (midpoints). This is far greater than than 2 odd vertices, so an Euler path cannot exist in this graph, and we cannot walk it traversing each edge exactly once.

Answer

The left picture can be drawn in one stroke of pen, without traversing a line twice. The right picture *cannot* be drawn in one stroke of pen, without traversing a line twice.

Problem 4

Does there exist a graph with 5 vertices which have the following degrees: 2, 4, 4, 4, 4?

Solution

Since the number of odd vertices is zero, the Handshaking Lemma is not violated, so the graph can still exist. Let Deg be a set of all vertex degrees: $Deg = \{2, 4, 4, 4, 4\}$. As per [4, p. 35], for a graph $\mathcal{G}(V, E)$ the number of edges is equal to the half of sum of all vertex degrees:

$$m = |E(\mathcal{G})| = \frac{1}{2} \sum_{i=1}^{|Deg|} deg_i. \tag{4}$$

In our case $m = (2 + 4 + 4 + 4 + 4)/2 = 18/2 = 9$. In other words, we need to check if a graph $\mathcal{G}(5, 9)$ exists with its vertex degrees $\in Deg$.

As per [1], there is only one isomorphism for a graph on 5 vertices with 9 edges – see Figure 5. The edges are enumerated and each vertex degree is denoted. We can easily see that the set of vertex degrees for this graph is $Deg' = \{3, 3, 4, 4, 4\}$, i.e. $Deg \neq Deg'$, so a graph $\mathcal{G}(5, 9)$ with its vertex degrees $\in Deg$ cannot exist.

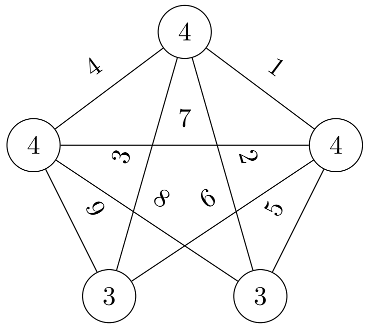


Figure 5: Only possible graph on 5 vertices with 9 edges

Answer

Such graph does not exist.

Problem 5

A connected graph on 10 vertices has 15 edges. What is the maximal number of edges one can remove so that the graph remains connected? Note that if your answer is N, then you need to explain that:

a) after removing some N edges from any connected graph on 10 vertices with 15 edges the resulting graph remains connected;

b) there exists a connected graph on 10 vertices with 15 edges such that after removing any $(N + 1)$ edges it becomes disconnected.

Solution

Let n be the number of vertices and m be the number of edges in the graph $\mathcal{G}(V, E)$. This graph cannot be a tree, since if it was a tree, it would have $m' = n - 1 = 10 - 1 = 9$ edges but we have $m = 15$. It means that $\mathcal{G}(10, 15)$ has at least one cycle. The simplest possible connected graph is a tree, so we need to break all cycles in our graph to make it a tree. As per [2, pp. 46–56], the circuit rank of a graph is the minimum number of edges which should be removed from the graph to break all its cycles, i.e., to make it a tree or forest, and is given by

$$r = m - n + c, \tag{5}$$

where r is the circuit rank, m is the number of edges, n is the number of vertices, and c is the number of connected components.

a) Since $\mathcal{G}(10, 15)$ is connected, $c = 1$, and by (5) $r = 15 - 10 + 1 = 6$. It means that we can safely remove at most some 6 edges to break all cycles and still keep the graph connected, now the graph will be a tree.

b) Following the result obtained in a), we can remove maximum 6 edges to still keep the graph connected. If we try to remove $r' = 6 + 1 = 7$ edges, we will immediately disconnect the graph. Indeed, if we rearrange (5) for c, we will get $c' = r' - m + n = 7 - 15 + 10 = 2$. It means that the graph will have 2 connected components, i.e. it will become disconnected.

Answer

The maximal number of edges one can remove so that the graph remains connected is 6. If we remove any 7 edges, the graph will become disconnected.

Problem 6

A graph on 6 vertices has 11 edges. Prove that this graph is connected.

Solution

The simplest possible connected graph is a tree with $n = 6$ vertices and $m = n-1 = 6-1 = 5$ edges. However, since the structure of the graph is not stated in the Problem, to guarantee connectedness of a graph with unknown structure, (3) must be true. In our case $n = 6, m = 11$. Let's check if m satisfies (3): $m' \geq (6-1) \cdot (6-2)/2 = 5 \cdot 4/2 = 10$, so m' must be at least 10. We have exactly $m = 11$, so the graph is guaranteed to be connected.

Another Approach

Let's find the maximal number of edges in a graph depending on the number of connected components in it. The maximal number of edges is the number of combinations of all pairs of connected vertices, i.e. the number of unique edges between every pair of distinct vertices (complete graph). We have a sequence of unordered pairs of connected vertices (edges) without repetitions of vertices, and the number of such 2-combinations is given by $\binom{n}{2}$, where n is the number of vertices. $\binom{n}{2} = \frac{n!}{2! \cdot (n-2)!} = \frac{n \cdot (n-1)}{2}$.

The number of possible connected components is the number of partitions of the number of vertices, i.e. the number of unordered ways to write the total number of vertices as a sum of positive integer numbers, and each summand in such a sum will represent the number of vertices in the corresponding connected component. In our case $n = 6$, so there 11 ways to partition the number 6. Let's list the maximal number of edges for each connected component (partition) in descending order:

- 6 components
  1. $6 = 1+1+1+1+1+1 \implies 6 \cdot \binom{1}{2} = 6 \cdot 0 = 0$ edges max
- 5 components
  2. $6 = 1+1+1+1+2 \implies 4 \cdot \binom{1}{2} + \binom{2}{2} = 4 \cdot 0 + 1 = 1$ edge max
- 4 components
  3. $6 = 1+1+2+2 \implies 2 \cdot \binom{1}{2} + 2 \cdot \binom{2}{2} = 2 \cdot 0 + 2 \cdot 1 = 2$ edges max
  4. $6 = 1+1+1+3 \implies 3 \cdot \binom{1}{2} + \binom{3}{2} = 3 \cdot 0 + 3 = 3$ edges max
- 3 components
  5. $6 = 2+2+2 \implies 3 \cdot \binom{2}{2} = 3 \cdot 1 = 3$ edges max
  6. $6 = 1+2+3 \implies \binom{1}{2} + \binom{2}{2} + \binom{3}{2} = 0+1+3 = 4$ edges max
  7. $6 = 1+1+4 \implies 2 \cdot \binom{1}{2} + \binom{4}{2} = 2 \cdot 0 + 6 = 6$ edges max
- 2 components
  8. $6 = 3+3 \implies 2 \cdot \binom{3}{2} = 2 \cdot 3 = 6$ edges max
  9. $6 = 2+4 \implies \binom{2}{2} + \binom{4}{2} = 1 + 6 = 7$ edges max
  10. $6 = 1+5 \implies \binom{1}{2} + \binom{5}{2} = 0 + 10 = 10$ edges max
- Finally, 1 component (connected graph)
  11. $6 = 6 \implies \binom{6}{2} = 15$ edges max

We can easily see that as the number of connected components increases, the maximal possible number of edges decreases. So even if the graph has just 2 components, it can have 10 edges at most, and the rest cases with more components have even fewer maximal possible number of edges. Thus, since our graph has 11 edges, it cannot have 2 or more connected components, i.e. it has to be connected.

Answer

The graph is connected. $m = 11 \in [m_1 = 5, m_2 = 15]$, where m_1 is for tree and m_2 is for complete graph.

Problem 7

A graph on 10 vertices has 3 isolated vertices (degree 0) and 7 vertices of degree 2. Could such a graph be bipartite? How many vertices are there in an optimal vertex cover for this graph? (Consider all possible cases.)

Solution

As per [5, p. 77], a graph can be bipartite, i.e. colored using no more than two colors, if it does not have odd cycles, i.e. closed paths with odd number of vertices. As for the 3 isolated vertices – let's denote each of them as $\mathcal{G}(1,0)$ – it is clear that they do not have any cycles and thus can be colored using two colors, in fact even one color is sufficient since those vertices are in different connected components. So, this part of the graph can be bipartite.

Now, let's consider the part with $n = 7$ vertices – let's denote it as $\mathcal{H}(V, E)$. As per (4), the number of edges $m = |E(\mathcal{H})| = 7 \cdot 2/2 = 7$. This graph cannot be a tree, since if it was a tree, it would have $m' = n - 1 = 7 - 1 = 6$ edges but we have $m = 7$. It means that $\mathcal{H}(7, 7)$ has at least one cycle.

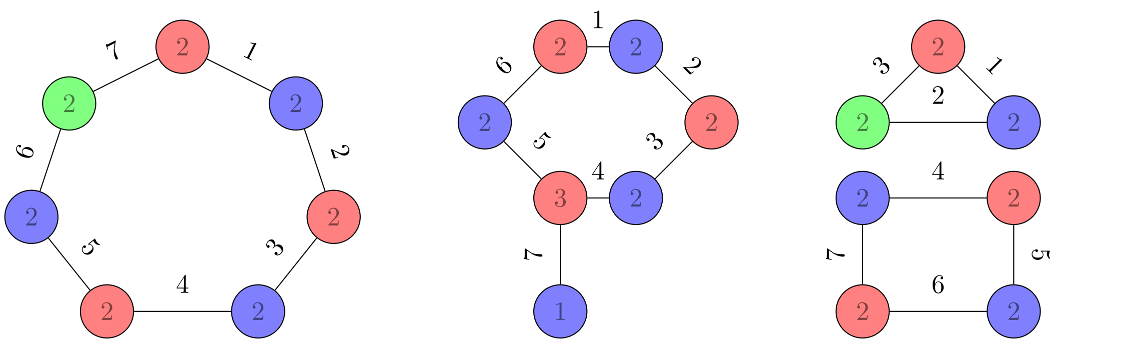
Applying the same approach as in Problem 6, we can list all partitions of the number of edges which is 7. However, this time we can only consider those partitions that satisfy all of the following conditions, regardless if the graph is connected or disconnected:

a) The number of edges $m = 7$, since we found that the 7-vertex part of the graph has 7 edges;

b) Each vertex degree is 2;

c) There are no odd cycles.

Out of 15 partitions of the number 7, there is no one that satisfies all of the conditions above. Let's show three 7-vertex graphs that violate only one condition above: 7, (6 + 1), and (3 + 4) – see Figure 6. The edges are enumerated and each vertex degree is denoted. Other cases are even worse violators.



1) connected
a) $m = 7$ (ok)
b) $\forall v_i, deg(v_i) = 2$ (ok)
c) one odd cycle (not ok)

2) connected
a) $m = 7$ (ok)
b) $\forall v_i, deg(v_i) \neq 2$ (not ok)
c) no odd cycles (ok)

3) disconnected
a) $m = 7$ (ok)
b) $\forall v_i, deg(v_i) = 2$ (ok)
c) one odd cycle (not ok)

Figure 6: Attempt to 2-color $\mathcal{H}(7, 7)$ with minimal violations

Finally, even though each of three $\mathcal{G}(1, 0)$ can be bipartite, $\mathcal{H}(7, 7)$ is not 2-colorable, so the whole graph is not bipartite.

As per [7, p. 25], vertex cover of a graph $\mathcal{G}(V, E)$ is a subset $C \subseteq V$ such that for any edge at least one its endpoint belongs to C, and a vertex cover is optimal, if it contains the smallest possible number of vertices. For each of three $\mathcal{G}(1, 0)$ it is obvious that any (including optimal) vertex cover $C(\mathcal{G}(1, 0)) = \{\ \}$ and $|C(\mathcal{G}(1, 0))| = 0$, since those graphs have no edges. For $\mathcal{H}(7, 7)$ there are only two graphs with $\forall v_i, deg(v_i) = 2$ – they are shown as 1) and 3) in Figure 6. We can visually verify both 1) and 3), and see that the optimal vertex cover $C(\mathcal{H}(7, 7)) = \{ v_1, v_3, v_5, v_6 \}$ and $|C(\mathcal{H}(7, 7))| = 4$ – see Figure 7.

So, the size of the optimal vertex cover of the whole graph is $3 \cdot |C(\mathcal{G}(1, 0))| + |C(\mathcal{H}(7, 7))| = 3 \cdot 0 + 4 = 4$.

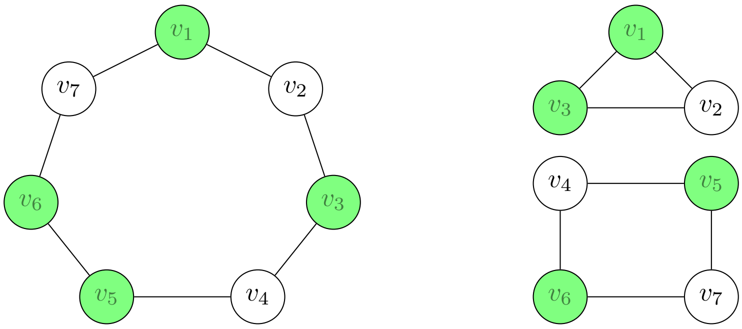


Figure 7: Optimal vertex covers of $\mathcal{H}(7, 7)$

Answer

This graph cannot be bipartite. Optimal vertex cover's size is 4.

References

[1] H.N. de Ridder et al. *Information System on Graph Classes and their Inclusions (ISGCI)*. URL: https://www.graphclasses.org/smallgraphs.html.

[2] Stepan Kuznetsov. *Connectedness*. Computer Science Department, Higher School of Economics. URL: https://edu.hse.ru/tokenpluginfile.php/152de50775e0367b2f7e6c5ffe5fbae4/2009467/mod_resource/content/1/03_connectedness.pdf.

[3] Stepan Kuznetsov. *Cycles in Graphs*. Computer Science Department, Higher School of Economics. URL: https://edu.hse.ru/tokenpluginfile.php/152de50775e0367b2f7e6c5ffe5fbae4/2009409/mod_resource/content/1/02_cycles.pdf.

[4] Stepan Kuznetsov. *Degrees and Distances*. Computer Science Department, Higher School of Economics. URL: https://edu.hse.ru/tokenpluginfile.php/152de50775e0367b2f7e6c5ffe5fbae4/2009451/mod_resource/content/1/01_degrees_and_distances.pdf.

[5] Stepan Kuznetsov. *Graphs*. Computer Science Department, Higher School of Economics. URL: https://edu.hse.ru/tokenpluginfile.php/152de50775e0367b2f7e6c5ffe5fbae4/2009397/mod_resource/content/1/01_graphs.pdf.

[6] Stepan Kuznetsov. *Random Graphs*. Computer Science Department, Higher School of Economics. URL: https://smartedu.hse.ru/mod/page/0/565454.

[7] Stepan Kuznetsov. *Subgraphs*. Computer Science Department, Higher School of Economics. URL: https://edu.hse.ru/tokenpluginfile.php/152de50775e0367b2f7e6c5ffe5fbae4/2009460/mod_resource/content/1/02_subgraphs.pdf.