Bioinformatics

Discrete Mathematics and Optimization. Final Exam (Retake): 13.12.2022

Write problems in separate sheets

Problem 1 (Combinatorics, 3 pt)**.** Answer the following questions. Each question can be answered independently of the rest

a) (1 pt) Compute the number of matrices of size $n \times m$ whose entries are $\{-1, 0, 1\}$.

b) (1 pt) Compute the coefficient of x^3y^2z in the expansion of $(x + 2y + z)^6$.

c) (1 pt) Prove by induction that $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$.

Solution.

a) There are a total of $n \cdot m$ entries on such a matrix, and each entry can be chosen independently among 3 possible values. Hence, we have 3^{nm}.

b) This coefficient is the multinomial coefficient with the corresponding power of 2: $2^2\binom{6}{3,2,1} = 2^2\frac{6!}{3!2!} = 2^2\frac{6\times5\times4}{2} = 240$.

c) The result clearly holds for $n = 1$. Assume the result to be true for n and let us check it for $n + 1$. We have then

$$1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2,$$

which is equal to

$$\frac{2n^3 + 3n^2 + n + 6n^2 + 12n + 6}{6} = \frac{2n^3 + 9n^2 + 13n + 6}{6},$$

and putting together the numerator one gets the expression

$$\frac{(n+1)(n+2)(2(n+1)+1)}{6}.$$

Problem 2 (Recurrences, 2 pt)**.** Obtain a close formula for a_n knowing that $a_n = a_{n-1} + 2a_{n-2}$ with $a_0 = 2$, $a_1 = 7$.

Solution. We look first for the characteristic polynomial, which is in our case $r^2 - r - 2 = (r-2)(r+1) = 0$, whose solutions are $r = 2$ and $r = -1$. So the general solution of the equation will be of the form:

$$a_n = A \cdot 2^n + B(-1)^n$$

By now fixing initial conditions:

$$n = 0 : A + B = 2,$$
$$n = 1 : 2A - B = 7.$$

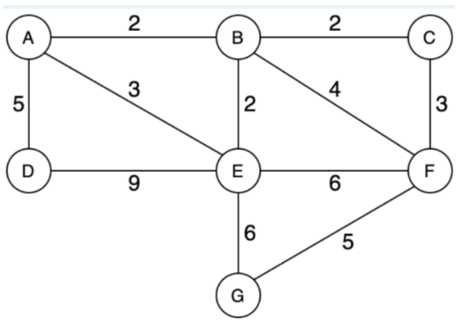
So we get that $3A = 9$, so $A = 3$, and $B = -1$. Hence, we have obtained that $a_n = 3\cdot2^n + (-1)^{n+1}$.

Problem 3 (Graphs, 3 pt)**.** Answer the following questions. Each question can be answered independently of the rest.

a) (1 pt) Define the degree of a vertex on a graph. Define the distance between two vertices on a graph. Finally, define the diameter of a graph.

b) (1 pt) Find the tree whose Prüfer code is $(2, 3, 3, 2, 1)$

c) (1 pt) Apply Kruskal algorithm to find the minimum spanning tree of the graph:



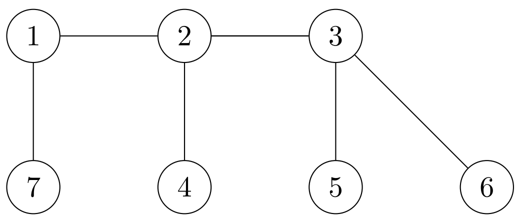
Solution.

a) For a graph $G = (V, E)$, the degree of $v \in V$ is the number of edges e in E such that $v \in e$. The distance between two vertices u and v (denoted by $d(u, v)$) is the minimal number of edges $e_1, e_2 \ldots e_r$ such that $u \in e_1$, $v \in e_r$ and that e_i and e_{i+1} share a vertex. The diameter of a graph G is the maximal possible distance between any pair of vertices in G.

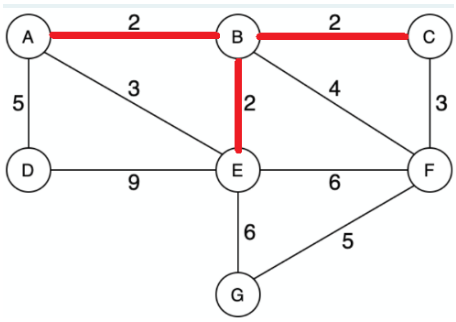
b) The algorithm to find the tree runs in 5+1 steps:

1) The smallest label not included in $(2, 3, 3, 2, 1)$ is 4: so we add edge $\{2, 4\}$ to $E = \emptyset$; we update the code with $(\mathbf{4}, 3, 3, 2, 1)$.
2) The smallest label not included in $(\mathbf{4}, 3, 3, 2, 1)$ is 5, so we add the edge $\{3, 5\}$ to $E = \{2, 4\}$; we update the code with $(\mathbf{4}, \mathbf{5}, 3, 2, 1)$.
3) The smallest label not included in $(\mathbf{4}, \mathbf{5}, 3, 2, 1)$ is 6, so we add the edge $\{3, 6\}$ to $E = \{\{2, 4\}, \{3, 5\}\}$; we update the code with $(\mathbf{4}, \mathbf{5}, \mathbf{6}, 2, 1)$.
4) The smallest label not included in $(\mathbf{4}, \mathbf{5}, \mathbf{6}, 2, 1)$ is 3, so we add the edge $\{2, 3\}$ to $E = \{\{2, 4\}, \{3, 5\}, \{3, 6\}\}$; we update the code with $(\mathbf{4}, \mathbf{5}, \mathbf{6}, \mathbf{3}, 1)$.
5) The smallest label not included in $(\mathbf{4}, \mathbf{5}, \mathbf{6}, \mathbf{3}, 1)$ is 2, so we add the edge $\{1, 2\}$ to $E = \{\{2, 4\}, \{3, 5\}, \{3, 6\}, \{2, 3\}\}$; we update the code with $(\mathbf{4}, \mathbf{5}, \mathbf{6}, \mathbf{3}, \mathbf{2})$.

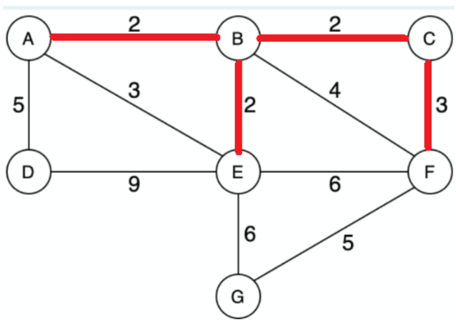
So, finally, as the labels not appearing in the final code are 1 and 7, we only need to add this edge, obtaining that the final set of edges is $E = \{\{2, 4\}, \{3, 5\}, \{3, 6\}, \{2, 3\}, \{1, 7\}, \{1, 2\}\}$. The resulting tree has the following form:



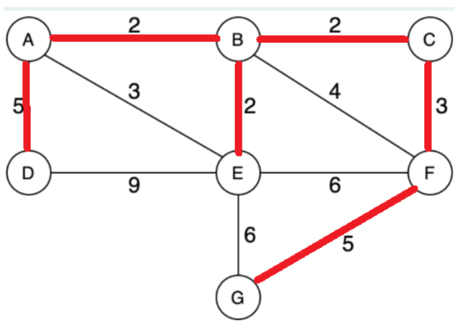
c) In Kruskal algorithm we order the edges according to their weights and we insert them (in order; ties in weights are ordered arbitrarily) as soon as we do not create a cycle. Once we order the edges, we observe that we can take all edges of weight 2:



After this step, only one edge of weight 3 can be taken:



Then, the edge of weight 4 cannot be taken. Finally, by taking the 2 edges of degree 5 we have taken 6 edges in total from a graph with 7 vertices, hence the spanning tree with minimum weight generated by Kruskal algorithm is the following:



Problem 4 (Linear Programming, 2pt)**.** We must run a program P and we have two available machines M_1 and M_2. Each execution of P spends $4s$ on M_1 and $1s$ on M_2. The cost per execution on M_1 is 1 cents and in M_2 is 4 cent. The energy generation by the heating of the machine per execution on M_1 is $2mW$ and on M_2 it is $1mW$. We must run at least 8 executions in total but we can not exceed $30s$ of total running time nor 20 cents of cost. We want to maximize the energy generation of the machines.

(a) (0.5 pt) Write a Linear Program to solve the problem.

(b) (0.5 pt) Draw the feasible region of the problem and identify a solution graphically.

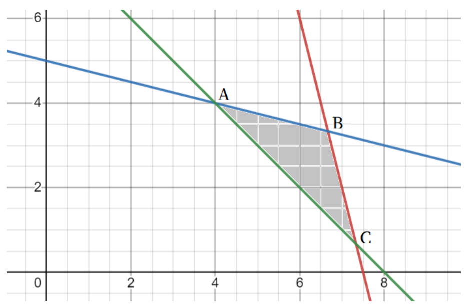
(c) (1 pt) Apply one iteration of the simplex algorithm by using variable x and starting at $(x, y) = (0, 0)$.

Solution.

(a) Let x and y be the number of executions of P on M_1 and M_2 respectively. The Linear Program is

$$
\begin{aligned}
\text{Maximize:} \quad & 2x + y \\
\text{Subject to:} \quad & 4x + y \leq 30 \\
& x + 4y \leq 20 \\
& x + y \geq 8 \\
& x \geq 0 \\
& y \geq 0
\end{aligned}
$$

(b) The feasible region is indicated in the next Figure, where $A = (4, 4)$, $B = (20/3, 10/3)$ and $C = (22/3, 2/3)$.



Hence, the solution to the problem must be reached at a border point. Just checking one gets that in A we have $8 + 4 = 12$, in B we get 16.3333 and in C we get $44/3 + 2/3 = 46/3 = 15.3333$. So the maximization is reached at point B.

(c) We introduce slack variables s_1, s_2, s_3 and write the problem as a maximization one by changing signs:

$$
\begin{aligned}
\text{Maximize:} \quad & 2x + y \\
\text{Subject to:} \quad & 4x + y + s_1 = 30 \\
& x + 4y + s_2 = 20 \\
& -x - y + s_3 = -8 \\
& x, y, s_1, s_2, s_3 \geq 0
\end{aligned}
$$

We start with the basic feasible solution $(0, 0, 30, 20, -8)$ on this problem and organize it in a tableau

$$
\begin{array}{llll|l}
s_1 & = 30 & -4x & -y & \\
s_2 & = 20 & -x & -4y & \\
s_3 & = -8 & +x & +y & \\
\hline
z & = 2x & +y & & z = 0
\end{array}
$$

Choose x as entering variable. The remaining variables put the restrictions $4x \leq 30$ from s_1, $x \leq 20$ form s_2 and $x \leq 8$ for s_3, the most restrictive is for s_1. We put $x = \frac{15}{2}$ and get the new solution $(15/2, 0, 0, 25/2, -1/2)$ and the new tableau

$$\begin{array}{llll|l}
x & = \frac{15}{2} & -\frac{1}{4}s_1 & -\frac{1}{4}y & \\
s_2 & = \frac{25}{2} & +\frac{1}{4}s_1 & -\frac{15}{3}y & \\
s_3 & = -\frac{1}{2} & -\frac{1}{4}s_1 & +\frac{3}{4}y & \\
\hline
z & = 15 & -\frac{1}{2}s_1 & +\frac{1}{2}y & z = 15
\end{array}$$

Observe that we need further iterations to get the optimal.