Albert-Ludwigs-Universität, Inst. für Informatik
Prof. Dr. Fabian Kuhn
M. Ahmadi, A. R. Molla, O. Saukh
January 13, 2016

Algorithm Theory, Winter Term 2015/16
Problem Set 9 - Sample Solution

Exercise 1: Maximum Matching (6 points)

We are given a bipartite graph $B = (U \dot{\cup} V, E)$ on two disjoint node sets U and V; each edge connects a node in U to a node in V. In the following, we define a *2-claw* to be a set of three distinct nodes $\{u, x, y\}$ such that $u \in U$, $x, y \in V$ and there is an edge from u to both nodes x and y.
We consider the following maximization problem on the graph B: Find a largest possible set of vertex-disjoint set of 2-claws. In other words, we want to find a largest possible subset of edges such that every node in U is incident to either 0 or 2 of the edges and each node in V is incident to either 0 or 1 of the edges (i.e., each node is either part of one 2-claw or it is not part of any 2-claw at all).

(a) (3 Points) Show that picking vertex-disjoint 2-claws in a greedy manner (as long as there is a 2-claw which is vertex-disjoint to all previously picked 2-claws, we pick it) results in a set of 2-claws which is at least one-third as large as an optimal set of vertex-disjoint 2-claws.

(b) (3 Points) In order to solve the problem optimally, let us now assume that in the given bipartite graph B, each node in U has at most 3 neighbors in V. Give a polynomial-time algorithm which computes a maximum set of vertex-disjoint 2-claws. You can use algorithms from the lecture as a subroutine.

Hint: Try to reduce the problem to the maximum matching problem in general graphs.

Solution:

(a) From the definition, each *2-claw* consists of 3 vertices and 2 edges. Any solution of the problem is a *vertex-disjoint* set of 2-claws. Let us assume that $\mathcal{S}_{opt}$ is an optimal solution for the graph B, and it consists of vertices in $U_{opt} \cup V_{opt}$ such that $U_{opt} \subseteq U$ and $V_{opt} \subseteq V$. It is clear that $|V_{opt}| = 2|U_{opt}|$, (from the definition of 2-claw).

Assume that $\mathcal{S}_{grd}$ is a greedy solution for the graph B and it consists of vertices in $U_{grd} \cup V_{grd}$ such that $U_{grd} \subseteq U$ and $V_{grd} \subseteq V$. For any 2-claw $\alpha = \{u, x, y\}$ in $\mathcal{S}_{opt}$, there are two possibilities:

1. Either α is in the greedy solution $\mathcal{S}_{grd}$, or
2. The greedy does not pick α because at least one of the three vertices $\{u, x, y\}$ of α is one of the vertices of a 2-claw which had been picked by the greedy algorithm beforehand. So we can claim that at least one-third of the vertices in $U_{opt} \cup V_{opt}$ are in $U_{grd} \cup V_{grd}$.

Considering the two possibilities, any time greedy picks a 2-claw, it misses to pick at most three 2-claw of the optimal solution. As a result, the number of 2-claws in $\mathcal{S}_{grd}$ is at least one-third of the number of 2-claws in $\mathcal{S}_{opt}$.

(b) To solve the problem, we reduce it to a maximum matching problem. Considering the given graph $B = (U\dot{\cup}V, E)$, we construct a graph $B' = (V', E')$ such that $V' := V$ and E' is explained as follows. For any two vertices $b, c \in V$ which are connected to the same vertex $a \in U$ in B, we connect b and c by an edge in B'. In other words, E' is a set of edges which connects any two vertices of V' in B' where the two corresponding vertices in V has a common neighbor in U in graph B. Now we show that the maximum matching in the general graph B' corresponds to a maximum set of vertex-disjoint 2-claws in the given bipartite graph B. To show the correctness of the reduction we need to prove the following claims.

Claim 1: The number of edges in the maximum matching in the graph B' is the maximum number vertex-disjoint of 2-claws in the graph B.

To prove Claim 1, we first prove the following two claims:

Claim 2.1: If there exist a set of k vertex-disjoint 2-claws in B, then there exist at least k vertex-disjoint edges in B'.
Proof. The claim follows directly from the construction of graph B' that we explained above. For any two vertices in V which have a common neighbour in U there exists an edge connecting the two corresponding vertices of V' in B'. Hence for any 2-claw in B there exist an edge in B'. Moreover, since any two 2-claws are vertex-disjoint, all these corresponding edges in B' are also vertex-disjoint (since, $V' = V$). Therefore, there exist at least k vertex-disjoint edges in B'. □

Claim 2.2: If there exist a set of k vertex-disjoint edges in B', then there exist at least k vertex-disjoint 2-claws in B.
Proof. We prove this claim based on the construction of the graph B' and also the fact that the degree of any vertex in U is at most 3. From the construction of B', the k vertex-disjoint edges in E' represent k distinct pairs of vertices in V such that the two vertices of any pair have a common neighbour in U. However, from the problem definition it is not possible to have two distinct pairs of vertices in V having a common vertex in U, since the degree of any vertex in U is at most 3. As a result, these k pairs of vertices belong to k vertex-disjoint 2-claws. □

Proof of Claim 1. Let us assume that $\mathcal{M}$ is the maximum matching in the graph B'. Based on Claim 2.2, we can conclude that there exist at least $|\mathcal{M}|$ vertex-disjoint 2-claws in B. But we also need to show that there does not exist a vertex-disjoint 2-claws set of size of more than $|\mathcal{M}|$.

Let us assume that there exist a set of vertex-disjoint 2-claws of size $|\mathcal{M}| + 1$ in the graph B. Then from the Claim 2.1, there must exist at least $\mathcal{M} + 1$ vertex-disjoint edges in B'. This contradicts our assumption that $\mathcal{M}$ is the maximum matching in B'. Hence the claim. □

We are now only left to show that how to compute a maximum matching in the general graph B'. For this we use the 'Edmond's Blossom Algorithm' (from the lecture notes) as a subroutine. The running time of Edmond's Blossom algorithm to compute the maximum matching in general graphs is $O(mn^2)$, which is polynomial.

Exercise 2: Triangles in Random Graphs (10 points)

Given a fixed vertex set $V = \{v_1, v_2, \ldots, v_n\}$ with n being an even number. Then the following (random) process defines the (undirected) random graph $G_p = (V, E_p)$:
For each vertex pair $\{v_i, v_j\}, i \neq j$ we independently decide with probability p whether the edge defined by this pair is part of the graph, i.e., whether $\{v_i, v_j\}$ is an element of the edge set E_p.
Furthermore we say that a subset $T = \{v_i, v_j, v_k\}$ of V of size 3 is a triangle of a graph, if all three edges $\{v_i, v_j\}$, $\{v_i, v_k\}$, $\{v_j, v_k\}$ are in the edge set of the graph.

(a) (1 Point) Let Z be the random variable that counts the number of edges in G_p. What kind of random variable is Z? What is the probability that Z has value k, for some k?

(b) (1 Point) Calculate m_T, the number of all triangles that could *possibly* occur in G_p.

(c) (2 Points) Let X denote the number of triangles in G_p. Calculate $\mathbb{E}[X]$.

The generation of the random graphs is now changed as follows. Before edges are determined each vertex is colored either red or green; we let K be the random variable that counts the number of red vertices. Between two red vertices there is an edge with probability p_{rr}, between two green vertices with probability p_{gg} and between vertices of different color with probability p_{rg} (edges are still picked independently).

(d) (3 Points) Assume first that with probability $\frac{1}{7}$ all vertices are red, with probability $\frac{2}{7}$ all vertices are green and with probability $\frac{4}{7}$ each vertex independently gets color red or green with probability $1/2$ each. Also $p_{rr} = 1$, $p_{rg} = \frac{1}{\sqrt{3}}$ and $p_{gg} = 0$. Calculate $\mathbb{E}[X]$ under these conditions!

(e) (3 Points) Let us now assume that K is not known, but it is known that $K \sim \text{Uniform}[1, n]$, i.e., in the painting process first the number of red vertices is determined and then K vertices are being selected to be red. The edge probabilities are the same as in the question above. Consider a vertex $v \in V$. Conditioned on the event that v is red, what is the probability that v is *not part of any triangle*?

Solution:

(a) To construct the n-node random graph, there are $\binom{n}{2}$ trials corresponding to $\binom{n}{2}$ pairs of vertices. In each trial the edge between one pair of nodes is picked independently with probability p. Hence, Z follows the binomial distribution with parameters $\binom{n}{2}$ and p; that is $Z \sim \mathit{Bin}\left(\binom{n}{2}, p\right)$.

Therefore, the probability of getting exactly k successes in $\binom{n}{2}$ trials is:

$$\Pr(Z = k) = \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k}$$

(b) Any set of 3 distinct nodes from the vertex set has non-zero probability to form a triangle in the graph. In other words, any set of 3 distinct nodes can possibly form a triangle. The number of these subsets is $\binom{n}{3}$. Hence, $m_T = \binom{n}{3}$.

(c) Any triple of (distinct) vertices in the graph can form a triangle. Consider all the $\binom{n}{3}$ triples of vertices in any order. Consider the random variable X_i for the i^{th} triple as follows:

$$X_i = \begin{cases} 1, & \text{if the } i^{th} \text{ triple forms a triangle} \\ 0, & \text{otherwise.} \end{cases}$$

Since the probability of any triple to form a triangle is p^3 (recall that probability of one specific edge being in E_p is p),

$$\mathbb{E}[X_i] = 1 \cdot \Pr(X_i = 1) + 0 \cdot \Pr(X_i = 0) = p^3$$

We can define the random variable X which represents the number of triangles in G_p, that is,

$$X = \sum_{i=1}^{\binom{n}{3}} X_i\,.$$

Therefore, using the linearity of expectation we have

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{\binom{n}{3}} X_i\right] = \sum_{i=1}^{\binom{n}{3}} \mathbb{E}[X_i] = \binom{n}{3} p^3\,.$$

(d) Let us define the following events:

$\mathcal{A}$: is the event when all the vertices are red.
$\mathcal{B}$: is the event when all the vertices are green.
$\mathcal{C}$: is the event when each vertex gets color red or green with probability 1/2.

$$\Pr(\mathcal{A}) = \frac{1}{7},\ \Pr(\mathcal{B}) = \frac{2}{7},\ \Pr(\mathcal{C}) = \frac{4}{7}$$

Then,

$$\begin{aligned}
\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{\binom{n}{3}} X_i\right] &= \sum_{i=1}^{\binom{n}{3}} \mathbb{E}[X_i] \qquad \text{[from the linearity of expectation]}\\
&= \sum_{i=1}^{\binom{n}{3}} \Pr(X_i = 1)\\
&= \sum_{i=1}^{\binom{n}{3}} \Big(\Pr(X_i = 1|\mathcal{A})\Pr(\mathcal{A}) + \Pr(X_i = 1|\mathcal{B})\Pr(\mathcal{B}) + \Pr(X_i = 1|\mathcal{C})\Pr(\mathcal{C})\Big) \quad \text{[law of total probability]}\\
&= \sum_{i=1}^{\binom{n}{3}} \Big(p_{rr}^3 \Pr(\mathcal{A}) + p_{gg}^3 \Pr(\mathcal{B}) + \Pr(X_i = 1|\mathcal{C})\Pr(\mathcal{C})\Big)\\
&\overset{**}{=} \sum_{i=1}^{\binom{n}{3}} \left(1\cdot\left(\frac{1}{7}\right) + 0\cdot\left(\frac{2}{7}\right) + \left(\frac{1}{8}\cdot 1^3 + \frac{1}{8}\cdot 0 + \frac{3}{8}\cdot 1\cdot\frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}} + \frac{3}{8}\cdot 0\cdot\frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\right)\cdot\frac{4}{7}\right)\\
&= \binom{n}{3}\cdot\left(\frac{1}{7} + \frac{2}{8}\cdot\frac{4}{7}\right) = \frac{2}{7}\cdot\binom{n}{3}
\end{aligned}$$

** For calculating $\Pr(X_i = 1|\mathcal{C})$ we apply the law of total probability on the color of the three nodes in the i^{th} triple of vertices. Note that the probability of having the three vertices all red or all green is 1/8 each, having two red vertices and one green vertex is 3/8, and one red vertex and two green vertices is 3/8.

(e) For calculating the probability of some vertex v to not be part of any triangle we apply the law of total probability on the nodes' color in the graph. For this purpose let us consider the following three possibilities for the value of random variable K.

(1) $K \geq 3$: In this case all the red nodes form a clique in the graph. Therefore, any red node is part of a triangle. Hence, the probability that any red node is not in a triangle is zero.

(2) $K = 1$: In this case the only red node cannot form a triangle with other nodes. The reason is that the probability of existence of an edge between any two green nodes is zero. Therefore, the probability that the red node is not part of a triangle is 1.

(3) $K = 2$: In this case the only possibility that the red node v is part of a triangle is that v forms a triangle with one green node and the other red node. Between the two red nodes there exists an edge with probability 1. Then, it is enough that there exist a green node which is connected to the both red nodes to have a triangle containing v. Therefore, the probability that each green node with the two red nodes does not form a triangle equals $\left(1 - 1\cdot\frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\right) = \frac{2}{3}$. Hence, since there are $n-2$ green nodes, the probability that v is not part of any triangle is $\left(\frac{2}{3}\right)^{(n-2)}$.

Due to the distribution of the random variable K, the probability of Case 1 is $\frac{n-2}{n}$, Case 2 is $\frac{1}{n}$, and Case 3 is $\frac{1}{n}$. Therefore, the probability p of one specific red node to not be part of any triangle, by applying law of total probability, is as follows:

$$p = \frac{n-2}{n} \cdot 0 + \frac{1}{n} \cdot 1 + \frac{1}{n} \cdot \left(\frac{2}{3}\right)^{(n-2)} = \frac{1}{n}\left(1 + \left(\frac{2}{3}\right)^{(n-2)}\right)$$