1. Graph Theory

Let $G$ be a connected simple graph that is not a cycle and is not complete. Prove that there exist distinct non-adjacent vertices $u, v \in V(G)$ such that the graph obtained from $G$ by deleting both $u$ and $v$ is connected.

**Solution:** Let us say that a vertex $v$ of a connected graph $H$ is *remote* if $v$ belongs to an end-block of $H$ and is not a cutvertex of $H$. (An end-block is a block that has degree at most one in the block decomposition tree of $H$.) Thus if $H$ is 2-connected, then every vertex of $H$ is remote. Furthermore, every connected graph on at least two vertices has at least two remote vertices.

If $G$ has a cutvertex $x$, then $G$ can be written as $G_1 \cup G_2$, where $V(G_1) \cap V(G_2) = \{x\}$. For $i = 1, 2$ let $z_i$ be a remote vertex in $G_i$, chosen in a block in which $x$ is not a remote vertex. Then $z_1, z_2$ are as desired.

Thus we may assume that $G$ is 2-connected. Since $G$ is not a cycle it has a vertex $u$ of degree at least three. Since $G$ is not complete the hypothesis implies that $u$ is not adjacent to some vertex $v \in V(G)$. Since we may assume that $G \setminus u \setminus v$ is disconnected, the graph $G \setminus u$ is not 2-connected, and hence has two non-adjacent remote vertices $a, b$ belonging to different end-blocks of $G \setminus u$. It follows that $G \setminus a \setminus b$ is connected, as desired.
2. Probability

Let \( \{X_n\} \) be a sequence of independent identically distributed random variables. Let

\[
S_n := X_1 + \cdots + X_n.
\]

Show that

\[
\frac{S_n}{\log n} \to 0 \quad \text{a.s.}
\]

implies that, for all \( c > 0 \), \( E e^{c|X_1|} < +\infty \).

Solution: The condition

\[
\frac{S_n}{\log n} \to 0 \quad \text{as } n \to \infty \quad \text{a.s.}
\]

implies that also

\[
\frac{S_{n-1}}{\log n} = \frac{S_{n-1}}{\log(n-1)} \frac{\log(n-1)}{\log n} \to 0 \quad \text{a.s.}
\]

Hence

\[
\frac{X_n}{\log n} \to 0 \quad \text{as } n \to \infty \quad \text{a.s.,}
\]

which means that, for all \( \varepsilon > 0 \), the following event

\[
\left\{ \frac{|X_n|}{\log n} \geq \varepsilon \text{ i.o.} \right\}
\]

has probability 0. The events

\[
\left\{ \frac{|X_n|}{\log n} \geq \varepsilon \right\}, \quad n = 1, 2, \ldots
\]

are independent. By the Borel-Cantelli Lemma, this implies that for all \( \varepsilon > 0 \)

\[
\sum_{n=1}^{\infty} \mathbb{P}\left\{ \frac{|X_n|}{\log n} \geq \varepsilon \right\} < +\infty.
\]

Since the random variables \( X_n \) are identically distributed, the last series can be also written as

\[
\sum_{n=1}^{\infty} \mathbb{P}\left\{ |X_1| \geq \varepsilon \log n \right\} < +\infty.
\]

Take \( \varepsilon = \frac{1}{e} \). Then we have

\[
\sum_{n=1}^{\infty} \mathbb{P}\left\{ e^{c|X_1|} \geq n \right\} = \sum_{n=1}^{\infty} \mathbb{P}\left\{ |X_1| \geq \varepsilon \log n \right\} < +\infty,
\]

which implies that \( E e^{c|X_1|} < +\infty \) (because for a nonnegative random variable \( Y \) we have \( EY < +\infty \) if and only if \( \sum_{n=1}^{\infty} \mathbb{P}\{Y \geq n\} < +\infty \).
3. Analysis of Algorithms

1. Let \( G = (V, E) \) be a graph and let \( w : E \to \mathbb{R}^+ \) be an assignment of nonnegative weights to its edges. For \( u,v \in V \) let \( f(u,v) \) denote the weight of a minimum \( u-v \) cut in \( G \). Show that for \( u,v,w \in V \),

\[
f(u,v) \geq \min\{f(u,w), f(w,v)\}.
\]

Generalize this to show that for \( u,v,w_1, \ldots, w_r \in V \),

\[
f(u,w) \geq \min\{f(u,w_1), f(w_1,w_2), \ldots, f(w_r,v)\}.
\]

2. Let \( T \) be a tree on a vertex set \( V \) with weight function \( w' \) on its edges. We will say that \( T \) is a flow equivalent tree if it satisfies the following condition: for each pair of vertices \( u,v \in V \), the weight of a minimum \( u-v \) cut in \( G \) is the same as that in \( T \). Let \( K \) be the complete graph on \( V \). Define the weight of each edge \( (u,v) \) in \( K \) to be \( f(u,v) \). Show that any maximum weight spanning tree in \( K \) is a flow equivalent tree for \( G \).

**Solution:**

1. Pick a minimum weight cut separating \( u \) and \( v \). It either separates \( u \) and \( w \), or \( w \) and \( v \). Thus, it is a \( u-w \) cut or a \( w-v \) cut, and the first part follows. Similarly, in the second part the minimum cut separating \( u \) and \( v \) separates \( w_i \) and \( w_{i+1} \) for some \( i = 0,1,\ldots,r+1 \), where \( w_0 \) means \( u \) and \( w_{r+1} \) means \( v \).

2. Let \( T \) be a maximum weight spanning tree of \( K \). Pick any two vertices \( u,v \in V \). The weight of a minimum \( u-v \) cut in \( G \) is \( f(u,v) \), by definition. The weight of a minimum \( u-v \) cut in \( T \) is the minimum weight edge in the unique path in \( T \) from \( u \) to \( v \). Let the path be \( w_0 = u, w_1, w_2, \ldots, w_r, w_{r+1} = v \). Note that the weights of the edges on this path are \( f(u,w_1), f(w_1,w_2), \ldots, f(w_r,v) \). Thus the minimum of these, by part (1), is no more than \( f(u,v) \). We deduce that equality holds, for if \( f(u,v) > f(w_i,w_{i+1}) \) for some \( i = 0,1,\ldots,r \), then by replacing in \( T \) the edge \( w_i, w_{i+1} \) by \( uv \) we obtain a spanning tree of strictly larger weight, contrary to the maximality of \( T \).
4. Linear Programming

I found a damaged sheet with the data of a linear programming program. This is what was on the sheet:

**Problem:**

\[
\begin{align*}
\text{minimize} & \quad 9x_1 - 2x_2 - 12x_3 + 31x_4 \\
\text{s.t.} & \\
1?x_1 - x_2 - 2x_3 + 2?x_4 & \geq 9 \\
-1?x_1 - x_2 - 1?x_3 + 2x_4 & \geq 10 \\
?x_1 + ??x_2 - ??x_3 - ?x_4 & \geq ? \\
-?x_1 + ??x_2 + ??x_3 - ?x_4 & \geq ? \\
??x_1 + ?x_2 + ?x_3 + ?x_4 & \geq -?? \\
\end{align*}
\]

\[x_1, x_2, x_3, x_4 \geq 0\]  

**Solution:**

\[< \text{computations}>\]

**Answer:** The optimal value is 1?.

Above, “?” stands for a decimal digit 0,1,...,9, perhaps different in different places. What is the optimal value in the problem? Justify your answer.

**Solution.** The problem was misstated. From the second constraint we get \(2x_4 \geq 10 + 1?x_1 + x_2 + 1?x_3 \geq 10 + 10x_1 + x_2 + 10x_3\), and hence \(9x_1 - 2x_2 - 12x_3 + 31x_4 \geq 9x_1 - 2x_2 - 12x_3 + 31(10 + 10x_1 + x_2 + 10x_3)/2 \geq 155\). Therefore, the optimal value is not of the form 1?.

\[\text{ACO Comprehensive Exam} \quad 19 \text{ March 2007}\]
5. Combinatorial Optimization

Given a set of positive numbers $b_1, \ldots, b_n$, consider the following mixed-integer set

$$S = \{(x, y) \in \mathbb{R}_+ \times \{0, 1\}^n : x + ay_i \geq b_i \quad i = 1, \ldots, n\},$$

where $a \geq \max\{b_i : i = 1, \ldots, n\}$. Consider a subset $R := \{i_1, \ldots, i_r\} \subseteq \{1, \ldots, n\}$ indexed such that $0 =: b_{i_0} < b_{i_1} \leq b_{i_2} \leq \cdots \leq b_{i_r}$, and the corresponding inequality

$$x + \sum_{k=1}^r (b_{i_k} - b_{i_{k-1}})y_{i_k} \geq b_{i_r}. \quad (1)$$

1. Prove that the above inequality (for any subset $R$) is valid for $\text{conv}(S)$.
2. Show that the above family of inequalities can be separated in polynomial time by solving an appropriate shortest path problem.

Solution: 1. Let $R = \{1, 2, \ldots, K\}$ such that $0 =: b_0 \leq b_1 \leq \cdots \leq b_K$. We prove by induction over $R$. For the base case $r = 1$, inequality (1) is the original inequality $x + ay_1 \geq b_1$ after tightening the coefficient of $y_1$ (since $a > b_1$). Suppose the inequality (1) corresponding to $\{1, \ldots, r\} \subset R$,

$$x + \sum_{k=1}^r (b_{i_k} - b_{i_{k-1}})y_{i_k} \geq b_r \quad (2)$$

is valid. We need to show that

$$x + \sum_{k=1}^{r+1} (b_{i_k} - b_{i_{k-1}})y_{i_k} \geq b_{r+1} \quad (3)$$

is valid. Note that the original inequality for $i = r + 1$ (after coefficient tightening)

$$x + b_{r+1}y_{r+1} \geq b_{r+1} \quad (4)$$

is valid. Then if $y_{r+1} = 0$ then (4) implies $x \geq b_{r+1}$ which in turn implies (3) since $\sum_{k=1}^{r+1} (b_{i_k} - b_{i_{k-1}})y_{i_k} \geq 0$.

If $y_{r+1} = 1$ then (3) reduces to (2) and so is valid.

2. We index $b_i$ such that $0 =: b_0 \leq b_1 \leq \cdots \leq b_n$. Given a solution $(x^*, y_1^*, \ldots, y_n^*)$ we need to find a set $R \subseteq \{1, \ldots, n\}$ for which (1) is violated. This can be done by solving a shortest path problem as follows. Construct a directed graph with nodes $\{s, 0, 1, \ldots, n, t\}$ such that there is an arc going from node $s$ to node 0 with length $x^*$, an arc going from each node $i \in \{1, \ldots, n\}$ to node $t$ with length $-b_i$, and an arc going from a node $i$ to a node $j$, for $i, j \in \{0, 1, \ldots, n\}$ with $i < j$, of length $(b_j - b_i)y_j^*$. Then it is easy to see that inequality (1) is violated if and only if there is an $s$-$t$ path with negative total length, and the nodes in $\{1, \ldots, n\}$ of such a path constitute the set $R$. 
6. Algebra

Let $G$ be a group of order 203. If $H$ is a normal subgroup of $G$ of order 7, then show that $H$ is contained in the center of $G$ and that $G$ is abelian.

**Solution:** Notice that $203 = 7 \cdot 29$. So if we consider that action of $G$ on itself by conjugation the orbits will have size 1 (if they are in the center), 7 or 29. Since $H$ is normal it is left invariant under conjugation. Thus the orbit of any element in $H$ under the action is contained in $H$. So the size of the orbit of any element in $H$ is either 1 or 7 and if one element has orbit size 7 then every element has orbit size 7. Since the identity is in $H$ and has orbit size 1, all the elements have orbit size 1. Thus $H$ is in the center of $G$.

Let $Z$ denote the center of $G$. We know $|Z| \geq 7$ and since $H$ is a subgroup of $Z$ the order of $Z$ is divisible by 7. Thus $|Z| = 7$ or 209. If the order is 209 the $G$ is abelian. So assume the order is 7. Thus $Z = H$ is a normal subgroup of $G$ and we can consider $G/Z$. This is a group of order 29 and hence must be cyclic. It is well know (or see below) that if $G/Z$ is cyclic then $G$ is abelian. This contradicts our assumption that $|Z| = 7$ and thus $|Z|$ must be 203 and $G$ is abelian.

(Suppose $G/Z$ is cyclic with generator $yZ$. So every element of $G/Z$ is of the form $(yZ)^n$ for some $n$. Thus every element of $G$ is of the form $y^n a$ for some $a \in Z$. Given two elements $g$ and $h$ in $G$ then write $g = y^n a$ and $h = y^m b$. We have $gh = y^n ay^m b = y^n y^m ab = y^m y^n ab = y^m by^n a = hg$. Here the second and fourth equality follow by $a, b \in Z$. So $G$ is abelian.)
7. Approximation Algorithms

Let $k$ be a power of two. Consider the following generalization of the Steiner forest problem to higher connectivity requirements: the specified connectivity requirement function $r$ maps pairs of vertices to $\{0, \ldots, k\}$, where $k$ is part of the input. Assume that multiple copies of any edge can be used; each copy of edge $e$ will cost $c(e)$. Give a factor $2 \cdot (\log_2 k + 1)$ algorithm for the problem of finding a minimum cost graph satisfying all connectivity requirements. You are allowed to use the Goemans-Williamson factor 2 Steiner forest algorithm as a subroutine.

Solution. Consider the connectivity requirements of all pairs of vertices written as $\lceil \log_2 k \rceil + 1$ bit integers. Consider the $i$th slice of these requirements as 0/1 requirement problem, i.e., a Steiner forest problem, for $0 \leq i \leq \lceil \log_2 k \rceil$. Solve it using Goemans-Williamson and multiply the solution by $2^i$. Take the union of these $\lceil \log_2 k \rceil + 1$ solutions as the final solution.

Let $OPT_f$ and $OPT$ be the optimal cost of the fractional and integral solutions to given problem. Let $OPT_i$ be the optimal cost of the fractional solution to the $i$th slice (as a 0/1 problem). Then, the cost of the solution produced is

$$2 \sum_{i=0}^{\lceil \log_2 k \rceil} 2^i OPT_i$$

Since $2^i OPT_i \leq OPT_f$, the claim follow.
7. Randomized Algorithms

Let $n$ be an odd number. There are $n$ cities $\{C_1, C_2, \ldots, C_n\}$ located at equidistant points on a circle. Two cars move at random from city to city, dropping packages at random in the visited cities. In particular, the process works as follows. At every time:

- Each car independently, chooses uniformly one of the two adjacent cities to the one where it is located and moves to it.

- With probability $p = n^{-4}$, the car in the city with the smallest number of packages dropped (among the two cities where the two cars are located) drops a new package. (Ties are broken randomly.)

1. Upper bound the number of packages in the city with the maximum number of packages. This should be a high probability bound (i.e., with probability tending to 1 when $n$ tends to infinity).

2. What would be the result if the $n$ cities were located at the vertices of a $d$-regular graph (instead of the cycle)? (Each car does a random walk on the vertices of the graph.)

Solution

1. Let $T_i$ the time when the $i^{th}$ package is dropped. Let $\mathcal{E}$ be the event $\{\forall i \leq n, T_{i+1} - T_i \geq n^{5/2}\}$. As $p = n^{-4}$, $\Pr[T_{i+1} - T_i < n^{5/2}] \leq n^{-3/2}$ and using the union bound, $\Pr[\neg \mathcal{E}] \leq n^{-1/2}$. The random walk on the cycle of length $n$ has mixing time $O(n^2 \log n)$. Therefore, conditioning on $\mathcal{E}$ the position of the cars at time $T_{i+1}$ is independent of the position at time $T_i$, and the process is equivalent to a bins and balls process where two bins are chosen independently at random and a ball is dropped to the least loaded bin. In this process the max load after dropping $n$ balls is $(1 + o(1))\frac{\log \log n}{\log 2}$.

2. Notice that the same argument shows that if the cities are located in a $d$-regular graph then the maximum load is $(1 + o(1))\frac{\log \log n}{\log 2}$ with probability $1 - 0(n^{-1/2})$. 