1. Computability, Complexity and Algorithms

Given an undirected graph $G = (V, E)$ with n vertices, two vertices $s, t \in V$ and an integer N, the #paths problem is to determine whether there exist at least N distinct s-t simple paths in G (note we say distinct, not disjoint). Use the following steps to show the problem is NP-complete by a reduction from the Hamitonian cycle problem.

1. Show that the number of simple cycles through a given edge of a given graph G can be counted by a reduction to the #paths problem.
2. Given a simple undirected graph $H = (U, F)$, let H' be obtained by subdividing each edge into ℓ edges and creating k parallel copies of each edge. Take a single Hamiltonian cycle in H. How many distinct Hamiltonian cycles in H' does it map to?
3. Suppose H is Hamiltonian and H' is constructed with $k = n, \ell = n+c$. Show that the total number of non-Hamiltonian simple cycles in H' is smaller than the number of Hamiltonian cycles in H' by a factor of n^c.
4. Show that the problem of deciding whether a given graph has a Hamiltonian cycle can be reduced to the #paths problem in polynomial time.

Solution:

1. For an edge (x, y), delete it from the graph, then count the number of paths from x to y using a binary search. This is the number of cycles through (x, y) in G.
2. Each edge is replaced by k^ℓ paths between the endpoints of the edge. So a single Hamiltonian cycle becomes $k^{\ell n}$ cycles.
3. A crude bound on the total number of simple cycles in H' of length at most $n-1$ is $k^{\ell(n-1)} \cdot n! \cdot 2^n$. The required condition is given by
$$k^{\ell n} \geq n^c \cdot k^{\ell(n-1)} n! 2^n$$
which is implied by $k^\ell \geq n^{c+n}$ and the statement follows.
4. Given a graph $G = (V, E)$, to determine if it is Hamitonian, first apply the subdivision and parallel copying of edges with $k = n, \ell = n+c$ to get a graph G'. Then apply the algorithm from Part (1) to count the number of simple cycles through each edge of G'. The sum of these counts is at most $n\ell$ times the number of simple cycles in G' (since each cycle can be counted at most as many times as the number of vertices in G'). Declare G is Hamiltonian iff this estimate of the number of cycles in G' is at least $k^{\ell n}$. If the original graph G is not Hamiltonian, then the total number of cycles in G' would smaller by a polynomial factor smaller, so even an n^{c-2} approximation for the #paths problem would suffice.

2. Theory of Linear Inequalities

Let x^* be a fractional extreme point of a rational polytopte $P := \{x \in \mathbb{R}^n \mid Ax \leq b\}$. Prove that there exists a Chvátal-Gomory cut for P that separates x^*.

Solution: Without loss of generality we may assume that A and b are integral. Since x^* is an extreme point of P, there exists constraints $\tilde{A}x \leq \tilde{b}$, a subsystem of $Ax \leq b$, such that

$$\{x \mid \tilde{A}x = \tilde{b}\} = \{x^*\}. \tag{1}$$

Also, without loss of generality we may assume that $\tilde{A}$ has full row rank. Therefore by Integer Farkas Lemma, there exists a rational vector y such that

$$y^\top \tilde{A} \in \mathbb{Z}^{1\times n},\ y^\top \tilde{b} \notin \mathbb{Z}.$$

Let z be an non-negative integral vector such that $z + y \geq 0$. Let $u = y + z$. Then by integrality of A and b we obtain that

$$\begin{aligned} u &\geq 0 \\ u^\top \tilde{A} &\in \mathbb{Z}^{1\times n} \\ u^\top \tilde{b} &\notin \mathbb{Z} \end{aligned} \tag{2}$$

Thus

$$u^\top \tilde{A}x \leq \lfloor u^\top \tilde{b} \rfloor$$

is a valid CG cut for P. Also (1) and (2) imply that this CG cut separates x^*.

3. Graph Theory

Let G be a 2-connected plane graph and let $V(G), E(G), F(G)$ denote its set of vertices, set of edges, set of faces, receptively. Let $\sigma : V(G) \cup F(G) \to \mathbb{Z}$ such that $\sigma(x) = d(x) - 4$ for all $x \in V(G) \cup F(G)$, where $d(x)$ is the number of edges incident with x. Show that

(1) $\sum_{x\in V(G)\cup F(G)} \sigma(x) = -8$.

(2) If $\delta(G) \geq 5$ then G contains K_4^- (obtained from K_4 by removing an edge) as a subgraph.

Solution: Using $\sum_{v\in V(G)} d(v) = 2|E(G)|$ and $\sum_{f\in F(G)} d(f) = 2|E(G)|$, (1) follows from Euler's formula. To prove (2), suppose $\delta(G) \geq 5$ and G contains no K_4^-. Note that $\sigma(x) < 0$ iff x is a face bounded by a triangle; in which case, $\sigma(x) = -1$. Modify σ as follows: For each vertex v and each triangular face f incident with v, subtract $1/3$ from $\sigma(v)$ and add $1/3$ to $\sigma(f)$. Let $\tau : V(G) \cup F(G) \to \mathbb{Z}$ denote the resulting function obtained from σ. Then $\sum_{x\in V(G)\cup F(G)} \tau(x) = \sum_{x\in V(G)\cup F(G)} \sigma(x)$. For each triangular face $f \in F(G)$, $\tau(f) = \sigma(f) + 1/3 + 1/3 + 1/3 \geq 0$. Since

G does not contain K_4^-, each $v \in V(G)$ is incident with at most $d(v)/2$ triangular faces; so

$$\tau(v) \geq d(v) - 4 - d(v)/6 = (5d(v) - 24)/6 \geq 1/6,$$

as $\delta(G) \geq 5$. Thus $\sum_{x \in V(G) \cup F(G)} \tau(x) > 0$, a contradiction as $\sum_{x \in V(G) \cup F(G)} \sigma(x) = -8$.

4. Algebra

(a) Let R be an integral domain containing a field k as a subring. Suppose that R is a finite dimensional vector space over k under the ring multiplication. Show that R is a field.

(b) Show that the conclusion does not hold without the assumption of being finite dimensional. That is, give an example of a ring R containing a field k as a subring, such that R is an integral domain but is not a field.

(c) Show that the conclusion does not hold without the assumption of being an integral domain. That is, give an example of a ring R containing a field k as a subring, such that R is finite dimensional over k but is not a field.

Solution:

(a) Let r be a nonzero element in R. Since R is finite dimensional over k, the set $\{1, r, r^2, \dots, r^n\}$ is linearly dependent over k for some n. Therefore there are $a_0, a_1, \dots, a_n \in k$, not all of them zero, such that $a_0 + a_1 r + a_2 r^2 + \cdots + a_n r^n = 0$. If $a_0 = 0$, then we would have $r(a_1 + a_2 r + \cdots + a_n r^{n-1}) = 0$. Since R is an integral domain, r is not a zero-divisor, so we would have $a_1 + a_2 r + \cdots + a_n r^{n-1} = 0$. By renaming the a_i's if needed, we can assume that $a_0 \neq 0$. Since a_0 is a nonzero element of the field k, it is invertible, so we have

$$1 = -a_0^{-1} r(a_1 + a_2 r + \cdots + a_n r^{n-1}).$$

This shows that r has an inverse in R for every nonzero $r \in R$, so R is a field.

(b) Let $R = \mathbb{Q}[x]$. Then R is an infinite dimensional vector space over $\mathbb{Q}$ and is an integral domain because the product of two non-zero polynomials is non-zero, which can be seen from the coefficients of highest degree terms. But R not a field since x does not have an inverse in R.

(c) Let $R = \mathbb{Q}[x]/\langle x^2 \rangle$. Then R is a two-dimensional vector space over $\mathbb{Q}$ but is not an integral domain because x is a zero-divisor.

4. Linear Algebra

Let A, B be $n \times n$ matrices. Show that $\sigma(AB) = \sigma(BA)$. (Recall the spectrum of A, $\sigma(A) = \{\lambda : A - \lambda I \text{ is not invertible}\}$.)

Solution: Suppose $\lambda \in \sigma(AB)$ and $\lambda \neq 0$. Let $u \neq 0$ and $ABu = \lambda u$. Then $BABu = \lambda Bu$. If $Bu \neq 0$, then $\lambda \in \sigma(BA)$. If $Bu = 0$, then we must have $0 = ABu = \lambda u$, which is a contradiction to $\lambda \neq 0$. By symmetry, we have shown that $\sigma(AB) \setminus \{0\} = \sigma(BA) \setminus \{0\}$.

Now suppose $0 \in \sigma(AB)$. If A is invertible, then B must be not invertible and so BA is not invertible. If A is not invertible, then it follows immediately that BA is not invertible. In either case we have $0 \in \sigma(BA)$. By symmetry, we have AB is not invertible if and only if BA is not invertible.

Thus $\sigma(AB) = \sigma(BA)$.

5. Analysis of Algorithms

Part a: Let $f(x)$ be a real-valued function. You are given a randomized algorithm B that, given input x and parameter $\epsilon > 0$, outputs $B(x)$ which approximates $f(x)$ as follows:

$$\forall x, \quad \mathbf{Pr}\left[(1-\varepsilon)f(x) \leq B(x) \leq (1+\varepsilon)f(x)\right] \geq 3/4. \tag{3}$$

Give an algorithm C that, given input x and parameters $\epsilon, \delta > 0$, outputs $C(x)$ satisfying:

$$\forall x, \quad \mathbf{Pr}\left[(1-\varepsilon)f(x) \leq C(x) \leq (1+\varepsilon)f(x)\right] \geq 1-\delta. \tag{4}$$

Achieve the best dependence on δ in $O(\cdot)$ notation (i.e., ignore constant factors).

Part b: Explain if your approach in (a) still works if instead of (3) the probability of success is weakened so that:

$$\forall x, \quad \mathbf{Pr}\left[(1-\varepsilon)f(x) \leq B(x) \leq (1+\varepsilon)f(x)\right] \geq 1/4. \tag{5}$$

Part c: Suppose that we have n polygons $P_1, \ldots, P_n$, all lying inside $[0,1] \times [0,1]$ which is the square with side length 1 on the Euclidean plane. Every polygon has area at least $\alpha > 0$. You are not given the polygons explicitly but instead for each polygon P_i we have access to a membership oracle: given a point $x \in \mathbf{R}^2$, the oracle returns YES if $x \in P_i$ and NO if $x \notin P_i$.

Give a randomized algorithm that approximately estimates the area of the *union* of these polygons. Given $0 < \varepsilon < 1$, the output Y of your algorithm should satisfy

$$\mathbf{Pr}\left[(1-\varepsilon)|P| \leq Y \leq (1+\varepsilon)|P|\right] \geq \frac{3}{4}$$

where $|P|$ denotes the area of $P = P_1 \cup \cdots \cup P_n$.

(You may assume that sampling a real number uniformly at random from $[0,1]$ takes constant time, and that each oracle call takes constant time.)

The running time of your algorithm should be polynomial in $n, 1/\varepsilon$, and $1/\alpha$.

Solution:

Part a: Given desired accuracy $\varepsilon > 0$ and input x, let $b_1, \ldots, b_N$ be independent samples from algorithm $B(x)$. Let $C(x)$ be the median of $b_1, \ldots, b_N$. For each $1 \le i \le N$, let

$$Z_i = \begin{cases} 1 & \text{if } b_i \in (1 \pm \varepsilon) f(x) \\ 0 & \text{otherwise.} \end{cases}$$

Let $Z = \sum_{i=1}^{N} Z_i$.

Note, $\mathbf{E}[Z_i] = \mathbf{Pr}[b_i \in (1 \pm \varepsilon) f(x)] \ge 3/4$, and hence

$$\mathbf{E}[Z] \ge \frac{3}{4} N.$$

Note if $Z > N/2$ then $C(x) \in (1 \pm \varepsilon) f(x)$. By Chernoff bounds,

$$\mathbf{Pr}[Z \le N/2] \le \mathbf{Pr}\left[|Z - \mathbf{E}[Z]| \le \frac{1}{4}\mathbf{E}[Z]\right] \le 2\exp(-\mathbf{E}[Z]/48) \le 2\exp(-N/64) \le \delta,$$

for $N \ge 64 \ln(2/\delta)$, and hence $N = O(\log(1/\delta))$ suffices.

Part b: We cannot boost the success probability. Consider the constant functions $f(x) = 1$ and $g(x) = 5$. Suppose for any input x the algorithm $B(x)$ outputs:

$$B(x) = \begin{cases} 1 & \text{with probability } 1/2 \\ 5 & \text{with probability } 1/2. \end{cases}$$

Then B satisfies (5) for both f and g. However, there is no algorithm C using B as a subroutine that satisfies (4) for both f and g.

Part c: The algorithm is as follows: in the k'th round, we choose a point x_k uniformly at random from $[0,1] \times [0,1]$ and we then check if the point lies in P_j for some j. If $x_k \notin P_j$ for all j, then let $Y_k = 0$; otherwise let $Y_k = 1$. The output of the algorithm is

$$Y = \frac{1}{N} \sum_{k=1}^{N} Y_k$$

where N is the number of rounds.

$$\mathbf{E}[Y] = \mathbf{E}[Y_k] = \mathbf{Pr}[Y_k = 1] = \mathbf{Pr}[x_k \in P] = |P|.$$

Then by Chernoff bounds,

$$\begin{aligned} \Pr(|Y - |P|| \ge \varepsilon |P|) &\le \Pr(|NY - N|P|| \ge \varepsilon N |P|) \\ &\le 2\exp\left(-N|P|\varepsilon^2/3\right) \\ &\le 2\exp\left(-N\alpha\varepsilon^2/3\right) \\ &\le 1/4, \end{aligned}$$

for $N = c/\alpha\varepsilon^2$ for sufficiently large constant c. Hence, $N = O(\alpha^{-1}\varepsilon^{-2})$ samples suffice, and the running time is $\mathrm{poly}(n, 1/\varepsilon, 1/\alpha)$.

6. Combinatorial Optimization

Recall that a graph G is *factor-critical* if for all $v \in V(G)$, $G - v$ has a perfect matching. An *open odd ear decomposition* of G is a sequence $H_0, H_1, \ldots, H_k$ of subgraphs of G such that, letting $G_j = \bigcup_{i=0}^{j} H_i$ for $j = 0, 1, \ldots, k$, we have

(a) G_0 is an odd cycle,

(b) for $i = 1, 2, \ldots, k$ the graph H_i is an odd length (i.e., odd number of edges) path with both (distinct) end vertices in $V(G_{i-1})$ and no internal vertex or edge in $V(G_{i-1})$, and

(c) $G = G_k$.

Show that if a 2-connected graph G is factor-critical, then it admits an open odd ear decomposition.

Solution. By Theorem 24.9 in [A. Schrijver, Combinatorial Optimization] the graph G has an odd cycle H_0 such that $G - V(H_0)$ has a perfect matching. Let us choose the maximum integer k such that there exist graphs $H_0, H_1, \ldots, H_k$ satisfying (a) and (b) such that there exists a perfect matching M in $G - V(G_k)$. Such a choice is possible, because $k = 0$ satisfies those requirements. If G has an edge $e \notin E(G_k)$ with both ends in G_k, letting H_{k+1} consist of e and its ends violates the maximality of k. We may therefore assume that no such edge exists and that there exists a vertex $u \in V(G) - V(G_k)$, for otherwise (c) holds. Since G is 2-connected, there exist two paths from u to G_k, vertex-disjoint, except for u. Thus there exists a path P of length at least two with both ends in G_k and otherwise disjoint from it.

Let us say that an edge $e \in M - E(P)$ is a *side edge* if at least one of its ends belongs to P. We may assume that P is chosen so that the number of its side edges is minimum. We may assume that there exists a side edge e, for otherwise letting $H_{k+1} := P$ contradicts the maximality of k. Let $e = uv$, where $u \in V(P)$. Let M' be a perfect matching in $G - u$. The subgraph of G with edge-set $M \triangle M'$ has a path Q with one end u. The other end of Q belongs to G_k.

By a *tail of* P we mean an M-alternating path with one end on P, the first edge a side edge and the other end in G_k. The argument of the previous paragraph shows that a tail exists. We may therefore choose a path P as above and a tail Q of minimum length. Let the ends of Q be $u \in V(P) - V(G_k)$ and $w \in V(G_k)$. If no internal vertex of P belongs to Q then the union of Q with one of the subpaths of P with end u (one that does not include w) contradicts the choice of P. We may therefore assume that there exists a vertex x that is internal to both P and Q and we may assume that x is chosen so that the subpath Q' of Q with ends u and x is as short as possible. The graph $P \cup Q'$ contains a unique cycle C. Let P' be obtained from P by replacing the subpath with ends u and x by Q. If the edge of $C \cap P$ incident with x does not belong to M, then P' contradicts the choice of P; otherwise the subpath Q'' of Q with ends x and w is a tail of P' and the pair P', Q'' contradicts the choice of P and Q.

Thus (c) holds, as desired.

7. Probabilistic methods

Prove that there is some constant $c > 0$ so that for every integer $k \geq 1$, given a graph and a set of k acceptable colors for each vertex such that every color is acceptable for at most ck neighbors of each vertex, there is always a *proper coloring* where every vertex is assigned one of its acceptable colors. (Recall as usual that a proper coloring requires that the endpoints of every edge get different colors.)

Solution: For each vertex, (independently) randomly assign it one of its acceptable colors. For each edge e and color i *that is acceptable for both endpoints*, consider the event $B_{e,i}$ that both endpoints of e are colored i.

Each bad event happens with probability at most $p = 1/k^2$. To construct the dependency graph, it suffices to connect each event to all other events whose corresponding edges share an endpoint. (Note that we only defined the events for edges whose endpoints have a common acceptable color.) In this dependency graph, each vertex has degree at most $O(ck^2) =: d$, since each endpoint has k acceptable colors and each of these colors is acceptable for at most ck neighbors.

If we ensure that $e(d+1)(1/k^2) < 1$, then we can apply the Local Lemma to deduce that with positive probability, all edges are properly colored. This shows that there is a suitable choice of $c < 1$ (less than 1/100 say), that guarantees the existence of such a coloring, since we may clearly assume that $k > 1$.