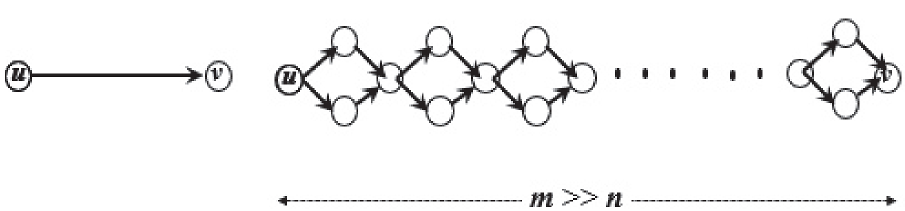
1. Computability, Complexity and Algorithms

(a): Count $s-t$ Paths in DAGs: Let $G(V,E)$ be a directed acyclic graph given in adjacency list representation, and let $s \in V$ and $t \in V$ be distinct vertices. Give an $O(|V|+|E|)$ algorithm that computes the number of distinct paths from s to t in G.
(b): Count $s-t$ Paths in General Directed Graphs: Let $G(V,E)$ be a general directed graph given in adjacency matrix representation, and let $s \in V$ and $t \in V$ be distinct vertices. Argue that, if there is a polynomial-time algorithm that computes the number of distinct paths from s to t in G, then there is a polynomial-time algorithm that decides Hamiltonicity in general directed graphs.

Solution: (a) First, in $O(|V|+|E|)$ time we perform topological sorting on G: That is, DFS with the vertex that finishes first being placed last, the vertex that finishes second being placed second to last etc. We we may now assume that the vertices of V are ordered $v_1, v_2, \ldots, v_n$ so that $v_i \to v_j \in E$ implies $i < j$. In addition, without loss of generality, we may assume that $s = v_1$ and $t = v_n$.
Next we count the number of distinct paths from v_1 to v_n using dynamic programming. Let us initialize $P(v_1) = 1$, and suppose that we have computed $P(v_{k'})$ for all $v_{k'}$ for $1 \le k' < k \le n$. We set $P(v_k) := \sum_{k' \to k \in E} P(v_{k'})$. In implementation:

Initialization: $P(v_1) = 1$
Memorization: for $k := 2$ to n
    $P(v_k) = 0$
    for all $v_k' \to v_k \in E$
        $P(v_k) := P(v_k) + P(v_{k'})$
Return: $P(v_n)$

Correctness follows from the recursive form of the solution, and running time from the fact that vertices and edges are used for updates a small constant number of times.
(b) Let $P(k)$ denote the number of paths of lenght k from s to t in G, $1 \le k \le n-1$. That is, each path contributing to $P(k)$ has $k+1$ vertices and k edges. Therefore, $P(k) < n^k$.
Define a graph $G'(V', E')$ by replacing each edge $u \to v$ with a chain of diamonds of length m, as below (m will be determined later.) Now a single edge of G gives rise to 2^m distinct paths from u to v. Now realize that $P(k) < n^k$ $s-t$ paths of length k of G give rise to at most



$\sum_{k=1}^{n-2} P(k) < \sum_{k=1}^{n-2} n^k = \frac{n^{n-1}-1}{n-1} - 1$ s–t paths of length at most $(n-2)$ in G, which is less than $\frac{n^{n-1}}{n-1} \times 2^{m(n-2)}$ s–t paths of length at most $2m(n-2)$ in G'.
On the other hand, a single s–t path of length $n-1$ in G gives $2^{m(n-1)}$ s–t paths of length

$2m(n-1)$ in G'.
By noting that $\frac{n^{n-1}}{n-1} \times 2^{m(n-2)} << 2^{m(n-1)}$, say for $m = n^2$

$$\begin{aligned}
\frac{n^{n-1}}{n-1} \times 2^{m(n-2)} &\quad \text{vs} \quad 2^{m(n-1)} \\
\frac{n^{n-1}}{n-1} &\quad \text{vs} \quad 2^{m} \\
\frac{n^{n-1}}{n-1} &\quad \text{vs} \quad 2^{n^2} \\
\frac{n^{n-1}}{n-1} &\quad << \quad 2^{n^2}.
\end{aligned}$$

Thus, if we can count $s-t$ paths in a general directed graph in polynomial time, then, we can can check for every pair $(s,t) \in V \times V$ if there is a length $n-1$ path, i.e. a Hamiltonian path, from s to t. And since there are exactly n^2 pairs (s,t), if we could count $s-t$ paths in a general directed graph, we could decide in polynomial time if there is a length $(n-1)$ directed path from s to t, we could decide Directed Hamiltonian Path, and hence Hamiltonicity, in polynomial time.

2. Analysis of Algorithms

Matrix Identity Testing

- Recall the Schwartz-Zippel lemma:

  Lemma 1 (Schwartz-Zippel Lemma) *Let $p(x_1, \ldots, x_n)$ be a nonzero polynomial of n variables with degree d. Let S be a finite subset of $\mathbb{R}$, with at least d elements in it. If we assign $x_1, \ldots, x_n$ values from S independently and uniformly at random, then*

  $$\mathbb{P}[p(x_1, \ldots, x_n) = 0] \leq \frac{d}{|S|}.$$

  Using the aforementioned lemma, design a randomized algorithm to test whether $AB = C$, where A, B, C are three $n \times n$ matrices. Analyze the probability with which it will succeed, and analyze its runtime.

- Explain how to "boost" the above algorithm to succeed with probability $1 - \delta$.

Solution:

- Pick a vector x uniformly from a set S^n. Calculate ABx and Cx. By n applications of the Schwartz-Zippel Lemma, the probability these are not equal if $AB \neq C$ is at least $1 - \frac{n}{|S|}$, and takes time $O(n^2)$.
  Alternatively one can argue using the principle of deferred decision without S-Z.

- Run the above algorithm $k = \left\lceil \frac{\ln(1/\delta)}{\ln(|S|/n)} \right\rceil$ rounds, and output "equal" if and only if all of the tests come back with the value of "equal".

3. Theory of Linear Inequalities

Let $e^k \in \mathbb{R}^n$ for $k = 0, \ldots, n-1$ denote the vector with the first k entries being 1 and the following $n-k$ entries being -1. Let $S = \{e^0, e^1, \ldots, e^{n-1}, -e^0, \ldots, -e^{n-1}\}$, i.e., S consists of all vectors consisting of $+1$ followed by -1 or vice versa.

1. Consider any vector $a \in \{-1, 0, 1\}^n$ such that
   (a) $\sum_{i=1}^n a_i = 1$, and
   (b) for all $j = 1, \ldots, n-1$, we have $0 \leq \sum_{i=1}^j a_i \leq 1$.

   Show that $\sum_{i=1}^n a_i x_i \leq 1$ and $\sum_{i=1}^n a_i x_i \geq -1$ are valid inequalities for $conv(S)$.
2. Show that any such inequality defines a facet of $conv(S)$.

Solution.

1. Pick any such $a \in \{-1, 0, 1\}^n$ satisfying the two conditions. Consider e^k for some $0 \leq k \leq n-1$. We have

$$\begin{aligned}
\sum_{i=1}^n e_i^k a_i &= \sum_{i=1}^k a_i - \sum_{i=k+1}^n a_i \\
&= 2\sum_{i=1}^k a_i - \sum_{i=1}^n a_i \\
&= 2\sum_{i=1}^k a_i - 1
\end{aligned}$$

   But since $\sum_{i=1}^k a_i \in \{0, 1\}$, we have $\left(2\sum_{i=1}^k a_i - 1\right) \in \{-1, 1\}$. Thus we have $-1 \leq a^T e^k \leq 1$ and equivalently, we have $-1 \leq a^T(-e^k) \leq 1$.

2. Consider again any such $a \in \{-1, 0, 1\}^n$ satisfying the two conditions. The above proof shows that for any $0 \leq k \leq n-1$, we have $a^T e^k \in \{+1, -1\}$. Thus either e^k or $-e^k$ satisfies the inequality at equality. In particular, we obtain n points in $conv(S)$ that satisfy it at

equality. We now show that these points are affinely independent to show it is a facet. Suppose there exist $\lambda_0, \dots, \lambda_{n-1}$ and $\epsilon_0, \dots, \epsilon_{n-1} \in \{-1, 1\}$ such that

$$\sum_{k=0}^{n-1} \lambda_k \epsilon_k e^k = 0,$$
$$\sum_{k=0}^{n-1} \lambda_k = 0.$$

We now show that all λ_i are zero by induction on i. Suppose that for some $i \in \{0, 1, \dots, n-2\}$ we have already shown that $\lambda_j = 0$ for $0 \leq j \leq i-1$. Now, we add the $(i+1)^{st}$ and last coordinate of each remaining e^k:

$$\sum_{k=i}^{n-1} \lambda_k \epsilon_k (e^k_{i+1} + e^k_n) = 2\lambda_i \epsilon_i (e^i_{i+1}),$$

where we use the fact that the $(i+1)^{st}$ and the last coordinate of e^k are of different signs except for e^i. Thus we obtain $\lambda_i = 0$.
For $i = n-1$, there is only a single vector remaining and thus λ_{n-1} must be zero.

4. Combinatorial Optimization

Assume n is odd, and $G = (V, E)$ is a graph with $|V| = n$, $|E| = 2n-2$, such that G is the union of two edge-disjoint spanning trees. Assume furthermore that half of the edges are colored red, the other half blue (where the coloring of edges is unrelated to the spanning trees). Show that G contains a spanning tree where exactly half of the edges are red and half of them blue.

Solution. Let R denote the edges colored red and B denote the edges colored blue. Let $\mathcal{M}_1 = (E, \mathcal{I})$ denote the graphic matroid and $\mathcal{M}_2 = (E, \mathcal{I})$ denote the partition matroid where a set $F \subseteq E$ is independent if there are at most $\frac{n-1}{2}$ red edges in F and at most $\frac{n-1}{2}$ blue edges in F. A common basis to the two matroids will be a spanning tree with half red and half blue edges. Thus it is enough to show that the common base polytope is non-empty. Indeed, we show $x^* = \frac{1}{2} \cdot 1_E$ is a feasible solution to the common base polytope. From Corollary 41.12d of Schrijver, it is enough to show that x^* is in the base polytopes of two matroids. Since G is a union of two edge disjoint spanning trees, say T_1 and T_2, we have $x^* = \frac{1}{2} 1_{T_1} + \frac{1}{2} 1_{T_2}$. Thus x^* is a convex combination of two bases of $\mathcal{M}_1$ and is in the base polytope of $\mathcal{M}_1$.

Partition the red edges in two arbitrary sets of size $\frac{n-1}{2}$, say R_1 and R_2 and similarly for the blue edges B_1 and B_2. Observe that $B_1 \cup R_1$ is a base of $\mathcal{M}_2$ and $B_2 \cup R_2$ is also a base. Since $x^* = \frac{1}{2} 1_{B_1 \cup R_1} + \frac{1}{2} 1_{B_2 \cup R_2}$, we have that x^* is in the base polytope of $\mathcal{M}_2$ as well. Thus the common base polytope is non-empty and any extreme point of it satisfies the conditions required.

5. Graph Theory

Let d be a positive integer and let G be a graph with average degree at least $8d$. Show that G contains a d-connected subgraph whose edges can be oriented so that the resulting digraph has no directed path on three vertices.

Solution: Note that the subgraph to be found must be bipartite. So we need to look for a d-connected bipartite subgraph in G.

By applying induction on $|V(G)|$, we can show that G contains a bipartite subgraph B such that $e(B) \geq e(G)/2$. ($e(G)$ denotes the number of edges in G.)

Thus, B has average degree at least $4d$. By a theorem of Mader (covered in MATH 6014), B has a d-connected subgraph, say H. Let V_1, V_2 be the bipartition of H, and orient all edges of H from V_1 to V_2.

6. Probabilistic methods

Let S_n be a random string of length n, where each character is, independently, chosen uniformly at random from the alphabet $\mathcal{A} := \{A, \ldots, Z\}$. For each n, let $H_n \in \mathcal{A}^m$ be a given string of length $m = m(n) \geq 0$. We say that S_n *contains* H_n if S_n it contains a consecutive substring of length m which equals H_n. Find a threshold function $m^* = m^*(n)$ such that

$$\Pr(S_n \text{ contains } H_n) \to \begin{cases} 1 & m = o(m^*), \\ 0 & m = \omega(m^*). \end{cases}$$

Solution: Let $X = X_n$ be the number of occurrences of H_n in T_n, and define E_i as the indicator for the event that T_n contains a consecutive copy of H_n starting at position i. Note that, for $1 \leq i \leq n - m + 1$, by independence of characters we have

$$\Pr(E_i) = (1/26)^m = 26^{-m}.$$

Hence

$$\mathbb{E}X = (n - m + 1)26^{-m},$$

so the natural guess for the threshold function (based on the heuristic $\mathbb{E}X \approx 1$) is, say,

$$m^* := \log_{26} n = \Theta(\log n).$$

For the 0-statement we use the first moment method: for $m = \omega(m^*)$ we have

$$\Pr(S_n \text{ contains } H_n) = \Pr(X \geq 1) \leq \mathbb{E}X \leq (n+1)26^{-m} \to 0,$$

For the 1-statement we shall use the second moment method. Clearly, for $m = o(m^*)$ we have

$$\mathbb{E}X = (n - m + 1)26^{-m} = \Theta(n) \cdot 26^{-m} \geq n^{1-o(1)} \to \infty.$$

Noting that sufficiently far apart strings are independent (as they depend on disjoint sets of independent random variables), it follows that

$$\mathrm{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \sum_{i,j\in[n-m+1]} [\Pr(E_i \cap E_j) - \Pr(E_i)\Pr(E_j)] \leq \sum_{i,j\in[n-m+1]:|i-j|\leq m} \Pr(E_i \cap E_j). \tag{1}$$

Estimating $\Pr(E_i \cap E_j) \leq \Pr(E_i) = 26^{-m}$, the crux is that for $m = o(m^*)$ we crudely have

$$\mathrm{var}X \leq O(n \cdot m) \cdot 2^{-m} \leq \Theta(\mathbb{E}X_n) \cdot o(\log n) = o((\mathbb{E}X_n)^2),$$

which implies the 1-statement using the second moment method (or Chebychev's inequality).

7. Algebra

Let p be a prime number. Show that if G is a finite p-group, and if $N \trianglelefteq G$ is a normal subgroup of order p, then N is contained in the center of G.

Solution: Consider the action of G on itself by conjugation. The orbit of any element of N is contained within N since $N \trianglelefteq G$. The size of the orbit of any element $x \in N$ is the index of the stabilizer of that element. By Lagrange's Theorem, the index of a subgroup of any group divides the order of the group, which is a power of p in this case. Thus, the size of the orbit is also a power of p. Since the orbit of x is contained in N and $|N| = p$, the orbit has size 1 or p. Our goal is to show that the orbit always has size 1, as this is the same as saying that every element of G centralizes x, i.e., that $x \in Z(G)$. Suppose then that the orbit had size p. But then G acts on N transitively. This is a contradiction since the identity is in its own conjugacy class.

Alternatively, note that the conjugation action of G acting on N gives a homomorphism from G into $\mathrm{Aut}(N)$, which is a group of size $p-1$. Thus, the image of this homomorphism is trivial, as it also has order a power of p, which implies that each element of N commutes with all elements of G.

7. Linear Algebra

Let A be a bistochastic matrix, that is a real $n \times n$ matrix such that

$$A_{i,j} \geq 0 \quad \forall i,j \qquad \sum_{i=1}^{n} A_{i,j} = 1 \quad \forall j \qquad \sum_{j=1}^{n} A_{i,j} = 1 \quad \forall i\ .$$

Let $a = \min_{i,j} A_{i,j}$ and let $v \in \mathbb{R}^n$ be such that $\sum_{i=1}^n v_i = 0$.

(a) Show that

$$\|Av\|_1 \leq (1 - na)\|v\|_1,$$

where $\|v\|_1 = \sum_{i=1}^n |v_i|$. Is the estimate sharp? That is, can you find A and v as above such that

$$\|Av\|_1 = (1 - na)\|v\|_1\ ?$$

(b) Show that

$$\|Av\|_\infty \leq (1-na)\|v\|_\infty,$$

where $\|v\|_\infty = \max_i |v_i|$. Is the estimate sharp? That is, can you find A and v as above such that

$$\|Av\|_\infty = (1-na)\|v\|_\infty \ ?$$

Solution: (a) Since $A_{i,j} \geq a$ and $\sum_{i=1}^n v_i = 0$ we have

$$\|Av\|_1 = \sum_i \left| \sum_j A_{i,j} v_j \right| = \sum_i \left| \sum_j (A_{i,j} - a)\, v_j \right| \leq \tag{2}$$

$$\sum_i \sum_j (A_{i,j} - a)\, |v_j| = (1-na)\|v\|_1 \ . \tag{3}$$

The estimate is sharp. Consider A such that $A_{i,i} = 1-(n-1)b$ while $A_{i,j} = b$ for $i \neq j$, where $b < 1/n$. Then $\min_{i,j} A_{i,j} = b$ and $Av = (1-nb)v$ for every v such that $\sum_{i=0}^n v_i = 0$.
(b) Call $I \subset \{1, \ldots, n\}$ the set of indices such that $v_i \geq 0$. Since $\sum_{i=0}^n v_i = 0$ both I and I^c are non empty. We have

$$\|Av\|_\infty = \max_i \left| \sum_j A_{i,j} v_j \right| = \max_i \left| \sum_{j\in I} A_{i,j} v_j - \sum_{j \in I^c} A_{i,j} |v_j| \right|$$

but we have

$$\sum_{j\in I^c} A_{i,j}|v_j| > a \sum_{j \in I^c} |v_j| = a \sum_{j\in I} v_j$$

so that

$$\sum_{j\in I} A_{i,j} v_j - \sum_{j\in I^c} A_{i,j}|v_j| \leq \sum_{j\in I}(A_{i,j} - a)v_j \leq \max_{j\in I} v_j \sum_{j\in I}(A_{i,j} - a) \leq (1-na) \max_{i \in I} v_i$$

in the same way

$$\sum_{j\in I} A_{i,j} v_j - \sum_{j\in I^c} A_{i,j}|v_j| \geq -(1-na) \max_{j\in I^c} |v_j|$$

so that

$$\|Av\|_\infty \leq (1-na) \max\left(\max_{i\in I} v_i,\ \max_{j\in I^c} |v_j| \right) = (1-na)\|v\|_\infty \ .$$

Clearly the same example as in part (a) works here to show that the estimate is sharp.