Computer Science and Engineering, IIT Palakkad
CS 5003: Parameterized Algorithms

The set $\{1, 2, \ldots, n\}$ is denoted by $[n]$. For an undirected graph G, let $V(G)$ and $E(G)$ denote the set of its vertices and the set of its edges, respectively. A path $P = (v_1, \ldots, v_\ell)$ is a sequence of distinct vertices such that $\{v_i, v_{i+1}\}$ is an edge for each $1 \leq i \leq \ell - 1$. The length of a path P is defined as the number of vertices in it and we call a path of length k as a k-path. The LONGEST PATH problem is defined as follows.

Input: An undirected graph G on n vertices and a positive integer k
Question: Does G have a path of length at least k?
Parameter: k

As HAMILTONIAN PATH is a special case of LONGEST PATH and is NP-hard, it follows that LONGEST PATH is NP-hard. Given an instance (G, k) of LONGEST PATH, we color the vertices of G uniformly at random from the color set $[k]$. Let $\chi : V(G) \to [k]$ denote this coloring. A path P is said to be χ-*colorful* if for any two distinct vertices $u, v \in V(P)$, $\chi(u) \neq \chi(v)$ holds. The dynamic programming algorithm that we discussed in the class leads to the following result.

Lemma 1. *Let G be an undirected graph and χ be a coloring on its vertices. If G has a χ-colorful k-path, then there is an algorithm that finds such a path in $\mathcal{O}(2^k n^2)$ time.*

Executing the algorithm given by Lemma 1 by iterating over all possible coloring functions χ (instead of random colorings) and declaring that G has no k-path if and only if all the executions fail to find a χ-colorful k-path, shows that LONGEST PATH can be solved in $\mathcal{O}(k^n 2^k n^2)$ time. In this lecture, we will discuss a smaller family of coloring functions that leads to a deterministic FPT algorithm for LONGEST PATH. We will prove the following theorem.

Theorem 2. *Given integers $n, k \geq 1$, there is a family $\mathcal{F}^*_{n,k^2}$ of coloring functions $\chi : [n] \to [k^2]$ of size $\mathcal{O}(n^2)$ that can be constructed in $n^{\mathcal{O}(1)}$ time satisfying the following property: for every set $S \subseteq [n]$ of size k, there is a function $\chi \in \mathcal{F}^*_{n,k^2}$ such that $\chi(u) \neq \chi(v)$ for any two distinct vertices $u, v \in S$.*

Then, we have the following result.

Theorem 3. LONGEST PATH *can be solved in $\mathcal{O}^*(2^k \binom{k^2}{k})$ time.*

Proof. Consider an instance $\mathcal{I} = (G, k)$ of LONGEST PATH. First, we compute the family $\mathcal{F}^*_{n,k^2}$ of $\mathcal{O}(n^2)$ coloring functions using Theorem 2 where n is the number of vertices in G. Then, for each coloring function $\chi : V(G) \to [k^2]$ in $\mathcal{F}^*_{n,k^2}$ and for each set $X \subseteq [k^2]$ of k colors, we determine if G has a χ-colorful k-path using colors from X. For this purpose, we delete the vertices of G that are colored using colors from $[k^2] \setminus X$ to obtain the graph G' and use Lemma 1 to determine if G' has a χ-colorful k-path or not. Due to the properties of $\mathcal{F}^*_{n,k^2}$ guaranteed by Theorem 2, it follows that $\mathcal{I}$ is a yes-instance if and only if there is a coloring function $\chi \in \mathcal{F}^*_{n,k^2}$ and a subset $X \subseteq [k^2]$ of k colors such that $G[\{v \in V(G) : \chi(v) \in X\}]$ has a χ-colorful k-path. The overall running time is $\mathcal{O}^*(2^k \binom{k^2}{k})$. $\square$

A faster algorithm for LONGEST PATH may be obtained by using the following variant of Theorem 2.

Theorem 4. *Given integers $n, k \geq 1$, there is a family $\widehat{\mathcal{F}}_{n,k}$ of coloring functions $\chi : [n] \to [k]$ of size $e^k k^{(\log k)} \log n$ that can be constructed in $e^k k^{(\log k)} n \log n$ time satisfying the following property: for every set $S \subseteq [n]$ of size k, there is a function $\chi \in \widehat{\mathcal{F}}_{n,k}$ such that $\chi(u) \neq \chi(v)$ for any two distinct vertices $u, v \in S$.*

Theorem 5. LONGEST PATH *can be solved in $\mathcal{O}^*((2e)^k k^{(\log k)})$ time.*

Proof. Consider an instance $\mathcal{I} = (G, k)$ of LONGEST PATH. First, we compute the family $\widehat{\mathcal{F}}_{n,k}$ of $e^k k^{\mathcal{O}(\log k)} \log n$ coloring functions using Theorem 4 where n is the number of vertices in G. Then, for each coloring function $\chi : V(G) \to [k]$ in $\widehat{\mathcal{F}}_{n,k}$, we determine if G has a χ-colorful k-path using Lemma 1. Due to the properties of $\widehat{\mathcal{F}}_{n,k}$ guaranteed by Theorem 4, it follows that $\mathcal{I}$ is a yes-instance if and only if G has a k-path that is colorful with respect to at least one of the coloring functions. The overall running time is $\mathcal{O}^*((2e)^k k^{(\log k)})$. $\square$

Proof of Theorem 2

Let U denote the universe $\{0, 1, \ldots, u-1\}$ where u is a prime number such that $n \le u \le 2n$. Let T denote $\{0, 1, \cdots, t-1\}$. For elements $a, b \in U$ with $a \ne 0$, let $h_{a,b} : U \to U$ be defined as $h_{a,b}(x) = (ax+b) \mod u$. Let $g : U \to T$ denote the function defined as $g(x) = x \mod t$. Let $\mathcal{F}_{u,t} = \{g(h_{a,b}) : a, b \in U, a \ne 0\}$. Observe that $|\mathcal{F}_{u,t}| = u(u-1) \le n^2$.

Lemma 6. *For each pair of distinct elements $x, y \in U$ and for each pair $a, b \in U$ with $a \ne 0$, we have $h_{a,b}(x) \ne h_{a,b}(y)$.*

Proof. Assume on the contrary that for some $a, b, x, y \in U$ satisfying the given properties, we have $h_{a,b}(x) = h_{a,b}(y)$. Then, $(ax+b) \equiv (ay+b) \mod u$ implying that $a(x-y)$ is divisible by u. As u is a prime, it follows that either a is divisible by u or $x-y$ is divisible by u. As $a \in U$ and $a \ne 0$, a cannot be divisible by u. As $x, y \in U$ and $x > y$, $x - y$ cannot be divisible by u. Hence, we arrive at a contradiction in each of the cases. $\square$

Lemma 7. *Given $q, r \in U$ such that $r, q \ne 0$, there is at most one element $p \in U$ satisfying $pq \equiv r \mod u$.*

Proof. Assume on the contrary that there are two distinct elements $p, p' \in U$ such that $pq \equiv r \mod u$ and $p'q \equiv r \mod u$. Without loss of generality assume that $p > p'$. Then, it follows that $(p-p')q$ is divisible by u. However, this leads to a contradiction as $r, q, p - p' > 0$ and $q, p, p' \in U$. $\square$

Lemma 8. *For every pair of distinct elements $x, y \in U$, $\Pr_{f \in_R \mathcal{F}_{u,t}}(f(x) = f(y)) \le \frac{1}{t}$.*

Proof. Consider a pair of distinct elements $x, y \in U$. Without loss of generality, assume $x > y$. We need to show that $\ell = |\{f : f \in \mathcal{F}_{u,t}, f(x) = f(y)\}| \le \frac{|\mathcal{F}_{u,t}|}{t}$. In order words, for any pair of distinct elements $x, y \in U$, there are at most $\frac{|\mathcal{F}_{u,t}|}{t}$ functions in $\mathcal{F}_{u,t}$ that map x and y to the same element in T. Observe that $\ell = |\{(a,b) : a, b \in U, a \ne 0, g(h_{a,b}(x) \mod u) = g(h_{a,b}(y) \mod u)\}|$.

Let $f \in_R \mathcal{F}_{u,t}$ where $f = g(h_{a,b})$. Let $(ax+b) \mod u = r$ and $(ay+b) \mod u = s$ for some $r, s \in U$. From Lemma 6, $r \ne s$. Consider the system of equations $ax + b \equiv r \mod u$ and $ay + b \equiv s \mod u$ where $r \ne s$. Then, $a(x-y) \equiv (r-s) \mod u$. As $x - y > 0$ and $r \ne s$, there is a unique solution for a in the range $\{1, \ldots, u-1\}$ from Lemma 7. Then, it follows there is a unique solution for b in the range $\{0, 1, \ldots, u-1\}$ as $b \equiv (r - ax) \mod u$. Consequently, there is exactly one pair (a, b) such that $a, b \in U$, $a \ne 0$, $(ax+b) \mod u = r$ and $(ay+b) \mod u = s$.

Thus, ℓ is upper bounded by the number of pairs $r, s \in U$ such that $r \ne s$ and $r \mod t = s \mod t$. For each choice of $r \in U$, there are at most $\lceil \frac{u}{t} \rceil - 1$ such choices for s. Then, it follows that $\ell \le u(\lceil \frac{u}{t} \rceil - 1) \le u(\frac{u+t-1}{t} - 1) = \frac{u(u-1)}{t} = \frac{|\mathcal{F}_{u,t}|}{t}$. $\square$

Lemma 9. *If $t = k^2$, then for each $S \subseteq U$ with $|S| = k$, there exists a function $f_S \in \mathcal{F}_{u,t}$ such that for each pair x, y of distinct elements of S, $f_S(x) \ne f_S(y)$.*

Proof. Consider a subset $S \subseteq U$ with $|S| = k$. Let $f \in_R \mathcal{F}_{u,t}$. Let $C(S)$ be the random variable denoting the number of pairs of distinct elements $x, y \in S$, such that $f(x) = f(y)$. For each pair of elements $x, y \in U$, define the random variable $Z(x,y)$ as $Z(x,y) = 1$ if $f(x) = f(y)$ and 0 otherwise. Then, $C(S) = \sum_{x,y \in S, x \ne y} Z(x,y)$. Now, $\mathsf{E}(C(S)) = \mathsf{E}(\sum_{x,y \in S, x \ne y} Z(x,y))$. By linearity of expectation, we have $\mathsf{E}(C(S)) = \sum_{x,y \in S, x \ne y} \mathsf{E}(Z(x,y)) = \sum_{x,y \in S, x \ne y} \Pr(f(x) = f(y))$. By Lemma 8, $\Pr_{f \in \mathcal{F}_{u,t}}(f(x) = f(y)) \le \frac{1}{t}$. Therefore, it follows that $\mathsf{E}(C(S)) \le \frac{k(k-1)}{2k^2} < 1$. Thus, there exists a function $f_S \in \mathcal{F}_{u,t}$ such that $C(S) = 0$. $\square$

Taking $\mathcal{F}^*_{n,k^2}$ as $\mathcal{F}_{u,k^2}$ restricted to the domain $[n]$ completes the proof of Theorem 2.

References

[1] N. Alon, R. Yuster, and U. Zwick. Color-Coding. *Journal of the ACM*, 42(4):844–856, 1995.

[2] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized Algorithms*. Springer, 2015.

[3] M. Mitzenmacher and E. Upfal. *Probability and Computing*. Cambridge University Press, 2017.

[4] M. Naor, L. J. Schulman, and A. Srinivasan. Splitters and Near-optimal Derandomization. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*, pages 182–191, 1995.