Fingerprinting

Idea: Instead of checking if $X = Y$, check if $h(x) = h(y)$ for some random-ish function $h$. If $h$ is not randomized, there is going to be worse case $x$ and $y$ that lead to false collision (show they are equal even if they are not). Pick $h$ such that selecting it is easier than checking for equality.

Matrix multiple checking

Determine if $A = B.C$ where, $A, B, C \in \mathbb{F}^{n \times n}$

Easy way: compute $B.C$ in $O(n^3)$ or $O(n^{2.36\ldots})$ using best matrix multiplication algorithm

Faster way: check if $Av = B.(Cv)$ for random vector $v$

Time $= O(n^2)$ because of just 3 matrix vector multiplications

If $A = B.C$ we get equality 100% of the time
But If $A \neq B.C$
Checking $(A - B.C)v = 0$ where $A - BC = X$ which is by assumption is not 0

Lemma: If $X \neq 0$, $Pr(X.v = 0) \leq \frac{1}{|\mathbb{F}|}$

Proof: If $X \neq 0$

$X$ has some $i^{th}$ row $u \neq 0$

$(Xv)_i = (u.v)$

Look at the special case where $X$ is a $1 \times n$ matrix

Bound $Pr(u.v = 0) = \frac{1}{|\mathbb{F}|}$

$u_j \neq 0$ for some $j$

$u.v = 0$ iff $u_jv_j = -\sum_{k \neq j} u_kv_k$

This happens iff

$v_j = -u_j^{-1}, \sum_{k \neq j} u_kv_k$

This says after we fix all but one coordinates of $v$, the last coordinate has to be exactly this one possibility. No matter what we fix this other coordinate, probability that $v_j$ is exactly this thing is one in the order of the field. Assume we have a finite field. For infinite field, we can’t take uniform vector over it.

This give us co-RP algorithm!
Polynomial Identity Testing

Polynomial $P(X_1, ...X_n)$ may be it is given by an arithmetic circuit. It is easy to compute $P$ given $X_1, ...X_n$, but it not easy to write down exactly what $P$ is in terms of its coefficients difficulty in expanding out large products and finding if there are any cancellations.

Question: is $P \equiv 0$? (identically zero, not at a point, but zero as a polynomial)

Easy randomized algorithm: Evaluate $P$ on random inputs. If $P$ not zero on any one input, then $P$ not identically zero.

Schwartz-Zippel (Thm 7.2)

Polynomial $Q(X_1, ...X_n) \in \mathbb{F}[X_i]$
Some set $S \subset \mathbb{F}$
  $r_i$ i.i.d. chosen uniformly from $S_i$
  if $Q \neq 0$
  $Pr(Q(r_1, r_2, ...r_n) = 0) \leq \frac{deg(Q)}{min|S_i|}$

Proof: Induction on $n$
  $n = 1$: $Q$ univariate polynomial of degree $d$
  $\Rightarrow$ if $Q \neq 0$, it has at most $d$ roots
  $\Rightarrow Pr(Q(r) = 0) = Pr(r \in T) \leq \frac{d}{|S|}$ where $|T| \leq d$

Assume its true for polynomials of $n - 1$ variables
Idea: $r_n \in S_n$
If $Q(X_1, ...X_n - 1, r_n) \neq 0$
  then probability $Q(r_1, ...r_n) = 0$ small

Let $d' =$ degree $Q$ in r.v. $X_1, ...X_{n-1}$
$Pr(Q(r_1, ...r_n) = 0) \leq Pr(Q(X_1, ...X_{n-1}, r_n) \equiv 0) + \frac{d'}{min|S_i|}$
Now we just need to bound the probability that $Q$ ends up being identically zero.

$Q(X_i, r_n) -$ coefficient of this are polynomial in $r_n$ is 0 iff all coefficients = 0
Coefficient of a term of degree $d' = poly(r_n)$ of degree at most $\leq d - d'$

By 1-variable case

$$Pr(coeff = 0) \leq \frac{d - d'}{|S_n|} \leq \frac{d'}{min|S_i|} + \frac{d - d'}{min|S_i|} = \frac{d}{min|S_i|}$$

Note:
Over infinite this is never a problem. Over finite fields, for example $P$ has degree 100 over $\mathbb{F}_2 \Rightarrow \mathbb{F}_{1024}$
So instead of thinking of $P$ as a polynomial over $\mathbb{F}_2$, consider it a polynomial over $\mathbb{F}_{1024}$
$P \equiv 0$ in $\mathbb{F}_2 \Leftrightarrow P \equiv 0$ in $\mathbb{F}_{1024}$
Take $S$ to be entire thing, $S = \mathbb{F}_{1024}$
$Pr(failure) < \frac{1}{10}$
Applications

Existence of perfect matchings

We have a bi-partite graph, edges are subset of $[n] \times [n]$.

$$E \subset [n] \times [n]$$

We want to know is there a permutation $\pi \in s_n$ s.t. $(i, \pi(i)) \in E$ for all $1 \leq i \leq n$

$$\det(A) = \sum_{\pi \in s_n} sgn(\pi) \prod A_{\pi(i)}$$

It has terms that test whether $A_{\pi(i)}$ is there or not, thus determinants are annoying. We have permutations, essentially the same without sign terms.

$$\text{Permanent}(A) = \sum_{\pi} \pi A_{\pi(i)}$$

If we consider $A_{ij}$

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

Then, Perm(A): number of perfect matchings. But it has some problems:

1. Computing permanent is \#P-hard (not something we can expect to do in polynomial time)
2. problem with det(A) is that it has cancellation - could get a false negative. For example, for a graph if we have matrix with all ones, we have determinant equal 0. Even though there is perfect matching hidden in there, since all terms are positive, they cancel each other.

To solve cancellation, make non zero terms variables instead of 1:

$$A_{ij} = \begin{cases} X_{ij} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

Claim: $\det(A)(X_{11}, \ldots, X_{nn}) \equiv 0$ iff no perfect matching

Algorithm: pick F elements or order $> n^2$

- Take $X_{ij}$ i.i.d. elements of \#
- Evaluate $\det(A)$ and check if 0
  - If it is non-zero, there has to be a perfect matching.

Color Edges

Perfect matching with exactly $k$ red edges To solve cancellation:

$$A_{ij} = \begin{cases} X_{ij} & \text{if } (i,j) \text{ not red} \\ X_{ij}Y & \text{if } (i,j) \text{ red} \\ 0 & \text{otherwise} \end{cases}$$

Is $\det(A)(X_{11}, \ldots, X_{nn}, Y)[Y^k] \equiv 0$