Concepts: Freivalds Technique + Fingerprinting, Schwartz-Zippel Thrm, Perfect matching in Graphs, Pattern matching (in strings).

Fingerprinting: Test equality of X and Y by choosing random mapping from $U \to V$ st. with high prob $X = Y$ iff. $image(x) = image(y)$ in V. $image(x)$ is fingerprint. Timecomplexity $\log(|V|)$ down from $\log(|U|)$.

Frevalds' technique: Gives $\mathcal{O}(n^2)$ randomized algorithm for verifying $AB = C$, without recomp. C and bounded error probability. Algorithm choose $r \in \{0,1\}^n$ at random, compute $x = Br, y = Ax = ABr$ and $z = Cr$. $AB = C \Rightarrow z = y$. Alg errors if $AB \neq C$ and $y = z$.

Thrm: Let A, B, C be $n \times n$ matrices over field $\mathcal{F}$ st. $AB \neq C$. Then for $r \in \{0,1\}^n$ chosen uniform. at random, $\mathbf{Pr}[ABr = Cr] \leq 1/2$.

Proof: Let $D = AB - C \Rightarrow D \neq \bar{0}$, bound prob. that $Dr = 0$. Assume 1'st row in D has non-zero entries and all non-zeroes preeced zerp entries.

$$k > 0.d = [v_1, v_2, .., v_k, 0, .., 0] = D_1$$

$\mathbf{Pr}[d^T r \neq 0]$ upper bound on $\mathbf{Pr}[Dr = 0]$ since first entry in Dr is $d^T r$. $d^T r = 0$ iff

$$r_1 = \frac{-\sum_{i=2}^{k} d_i r_i}{d_i}$$

Invoke Principle of Deffered Decisions, and assume $r_2, .., r_k$ is drawn before r_1. Then above rhs is fixed at $\exists v \in \mathcal{F}$. Since r_1 is uniform and ind. drawn from $\{0,1\}$

$$\mathbf{Pr}[v = r_1] \leq 1/2 \Leftrightarrow \mathbf{Pr}[ABr = Cr] \leq 1/2$$

.

Principle of Deffered Decisions: Assume that entire set of random choices is not made in advance, but rather at each step of the alg. we fix only the random choice that must be revelaed.

Poly compare $Q(x) = P(x)$ iff all coefficient agree. String represnted in fixed alphabet can be viewed as polynomial of degree n, k'th entry k'th coefficient.

Poly prod verfication problem: Given $P_1(x), P_2(x), P_3(x) \in \mathcal{F}_{[\S]}$ verify $P_1 P_2 = P_3$. P_1, P_2 deg $\leq n \Rightarrow$ P_3 deg $\leq 2n$. Eval. def n $\mathcal{O}(n)$, mult by Fast Fourier Trans $\mathcal{O}(n \log n)$.

Alg. for $S \subset \mathcal{F}$ $|S| \geq 2n + 1$. Pick $r \in S$ uni. at ran. Eval $P_1(r), P_2(r), P_3(r)$ declare correc tif $P_1(r)P_2(r) = P_3(r)$. Fails if $P_1(x)P_2(x) \neq P_3(x)$, but $P_1(r)P_2(r) = P_3(r)$.

$Q(x) = P_1(x)P_2(x) - P_3(x)$ of deg $\leq 2n$. Q is identically zero if all coefs are zero. If $Q(x) \equiv 0$, then with high prob. $Q(r) = 0$. $\leq 2n$ distinct roots $\Rightarrow \leq 2n$ distinct choices of $r \in S$ have $Q(r) = 0 \Rightarrow$ $\mathbf{Pr}[error] \leq 2n/|S|$.

Multivatirate poly $Q(x_1, .., x_n$ def of any term is sum of expons of the terms. Total deg of Q is max deg of terms.

Thrm (Schwartz-zippel) Let $Q(x_1, .., x_n) \in \mathcal{F}_{[x_1,..,x_n]}$ be a mult. var. poly of tot deg d. Fix $S \leq \mathcal{F}$ st. $|S| \in \mathcal{Z}^+$ and let $r_1, .., r_n$ be uni at ran from S. Then

$$\mathbf{Pr}[Q(r_1, .., r_n) = 0 | Q(x_1, .., x_n) \not\equiv 0] \leq \frac{d}{|S|}$$

Proof: By induction of num vars n. Base $n = 1$, $Q(x_1)$ of deg d, by Thrm 1 $\mathbf{Pr}[Q(r) = 0] \leq 1/|S| \leq d\|S|$. Assume inductive hypothesis $n > 1$ is true for $n - 1$ vars. Consider $Q(x_1, .., x_n)$ and factorize out x_1 to obtain

$$0 < k \leq d.Q(x_1, .., x_n) = \sum_{i=0}^{k} x_1^i Q_i(x_2, .., x_n)$$

for k largest exponent of x_1. x_1^k coeffs, $Q_k(x_2, .., x_n)$ is not ident. zero by choice of k. Tot deg $Q_k \leq d - k$ inductive hyp.

$$\Rightarrow \mathbf{Pr}[Q_k(r_2, .., r_n) = 0] \leq (d - k)/|S|$$

Suppose $Q_k(r_2, .., r_n) \neq 0$. Construct

$$q(x_1) = Q(x_1, r_2, .., r_n) = \sum_{i=0}^{k} x_1^i Q_i(r_2, .., r_n)$$

$q(x_1)$ has def k and is not ident. zero since coeffs of x_1^k is $Q_k(r_2,..,r_n)$. From basecase we have

$$\mathbf{Pr}[q(r_1) = Q(r_1,..,r_n) = 0] \leq k/|S|$$

So

$$\mathbf{Pr}[Q_k(r_2,..,r_n) = 0] \leq \frac{d-k}{|S|}$$
$$\mathbf{Pr}[Q(r_1,..,r_n)|Q_k(r_2,..,r_n)] \leq \frac{k}{|S|}$$

By $\mathbf{Pr}[\mathcal{E}_1 \leq \mathbf{Pr}[\mathcal{E}_1|\bar{\mathcal{E}}_2] + \mathbf{Pr}[\mathcal{E}_2]$

$$\mathbf{Pr}[Q(r_1,..,r_n) = 0] \leq \frac{k-k+d}{|S|}$$

Text X is a string of lenght $|T| = n$. Pattern Y, string of lenght $|Y| = m$ st $m \leq n$. Both over finite alph.
Pattern in X occurs if $\exists j \in \{1,..,n-m+1\}$ st for $1 \leq i \leq mX_{j+i-1} = Y_i$
Let $X(j) = x_j..x_{j+m-1}$ be substring of text X. A match at j $X(j) = Y$
Randomized algorithm: First match: choose fingerprint function to reduce cost of compare.

$$F_p(X(j)) = X(j) mod p$$

where $X(j)$ is m-bit int and p chosen uni. at ran. from primes $< \tau$.
Probability that alg errors.

$$\mathbf{Pr}[F_p(Y) = F_p(X(j))|Y \neq X(J)] \leq \frac{m}{\pi(\tau)} = \mathcal{O}\frac{m\log\tau}{\tau}$$

Must happen n times for alg to error. Choose $\tau = n^2 m \log n^2 m$ $\mathbf{Pr}[err] = \mathcal{O}(n^{-1})$.
Timecomplexity: For $1 \leq j \leq n-m+1$

$$X(j+1) = 2[X(j) - 2^{m-1}x_j] + x_{j+m} \Rightarrow$$
$$F_p(X(j+1)) = 2[F_p(X(j)) - 2^{m-1}x_j] + x_{j+m} \mod p$$

Gives $\mathcal{O}(n+m)$.
Las vegas version 1, expected runtime $\mathcal{O}((n+m)(1-1/n) + nm(1/n))$. Whenever a match occurs we check if with full compare. If this fails we run brute-force algorithm. High variance in runtime.
Las vegas version 2, restart if false match is detected and draw new p, prop of t restarts $\leq 1/n^t$. Low variance in runtime. Nontrival to find p.