Randomized Algorithms, Lecture No. 1
Russell Impagliazzo, Lecturer

Paul Tucker, faithful scribe

January 8, 1998

The big questions for this course are:

- Why make random decisions in general? (since we usually want a predictable answer)
- How much can it help us?

A theme for this course will be that there exists a randomness/ignorance tradeoff. Randomness can help us find values we know exist, but don't know what they look like. Conversely, sometimes non-computability can replace randomness.

Technique 1: A non-constructive counting argument. We have a general rule; is it always true? If not, then we can show it is frequently false by a counting argument. If not always true, then there must be an upper bound on the number of true instances. If the number of true instances is small, relative to the total number of instances, then straightforward testing on random instances can find a counter-example to the rule.

A simple example: testing Diophantine equalities.

$$17^{13256} + 19^{13256} \stackrel{?}{=} 21^{13256}$$

We can check this by testing rings: sampling at different congruencies. If true, then

$$\forall N \quad N | a_1 x_1^{e_1} + a_2 x_2^{e_2} + \cdots a_j x_j^{e_j} \quad = \quad 0. \tag{1}$$

How can we find an N that is not a factor, *i.e.* a counter-example?

Observe that if each a_i, x_i, e_i has at most n digits, and there are at most n terms total,

$$\begin{aligned} a_1 x_1^{e_1} + a_2 x_2^{e_2} + \cdots a_j x_j^{e_j} &= M && (2)\\ &\leq n(2^{n2^n}) && (3)\\ &\leq 2^{2^{2n}}. && (4)\end{aligned}$$

To figure out how many examples and counter-examples there are to the rule about congruence, we look at prime factorization.

$$M = p_1^{e_1} * p_2^{e^2} * \cdots * p_k^{e_k} \tag{5}$$

where

$$p_i \geq 2; M \geq 2^k.$$

Then the number of prime factors is at most 2^{2n}, since

$$k \leq \log M \leq 2^{2n}$$

The prime number theorem tells us that the number of primes $\leq x$ is approximately $\frac{x}{\ln x}$. The number of primes of length l (binary rep.) is $\Theta(\frac{2^l}{l})$. So, if we pick $l \geq 3n$, then the number of primes of length l is more than twice the number of prime factors of M.

Procedure:

1. Set $i \leftarrow 0$.
2. If $i \geq n^2$, output "looks OK".
3. Pick a random l-bit integer N, and set $i \leftarrow i + 1$.
4. Compute M mod N. (Using repeated squaring method.)
5. If the result is 0, repeat from step 2. If not, output "equation is false."

What's the justification for outputting "looks OK" after n^2 trials?

- If $M = 0$, we will never find a counter-example.
- If $M \neq 0$ then

$$\begin{aligned} \text{Prob}[N \text{ is prime } \& \ N \not| M] &= \text{Prob}[N \text{ is prime}] * \text{Prob}[N \not| M | N \text{ is prime}] \\ &\geq \text{Prob}[N \text{ is prime}] * 1/2 \\ &\geq \Theta(1/n) * 1/2 \end{aligned}$$

If $M \neq 0$ then the probability that our algorithm mistakenly outputs "looks OK" is

$$\begin{aligned} &\leq (1 - 1/cn)^{n^2} \\ &\leq ((1 - 1/cn)^{cn})^{n/c} \approx e^{-n/c} \\ &= \exp(-\Theta(n)). \end{aligned}$$

So, the probability of making a mistake is exponentially small, if sampling is independent and random.

Define RP to be the class of decision problems L with polynomial-time probabilistic algorithms $A(x,r)$:

$$\begin{aligned}\forall x,\quad x \in L \quad &\rightarrow \quad \Pr_r[A(x,r) = \text{Accept}] \geq 1/2 \\ x \notin L \quad &\rightarrow \quad \Pr_r[A(x,r) = \text{Reject}] = 1\end{aligned}$$

Observe that *invalid* Diophantine equations $\in$ RP; valid Diophantine equations $\in$ co-RP. The ability to prove or refute may not be symmetric.

More examples.

1. Given n, find a prime of length n. There is no known algorithm to solve this problem (with certainty) in polynomial time.

2. Given a number p, is it prime?

The first question reduces to the second, by the prime number theorem. If we pick random numbers, and test each for primality, we expect to pick only $O(n)$ before finding a prime.

The algorithm for testing primality, which we will describe next lecture, is based strongly on the following theorem:

A group is a set with a binary operator $*$, an identity element 1, and an inverse operator $^{-1}$, such that

- $x * x^{-1} = 1 = x^{-1} * x$
- $x * 1 = x = 1 * x$
- $(x * y) * z = x * (y * z)$.

Examples:

- addition mod N is a group on the numbers mod N, $0, ..N-1$.
- multiplication mod p is a group on the non-zero natural numbers (where p is prime) mod p, $1, ..p-1$.
- multiplication mod N is a group on the numbers mod N relatively prime to N , which is called Z_N^*.

The main step to seeing the latter two is to verify that inverses exist. If a is relatively prime to N, then $N|ax$ if and only if $N|x$ (all the prime factors for N must appear in x because none of them appear in a.) Moreover, if a and y are both rel. prime to N, then so is $aymodN$. So for each $a \in Z_N^*$ we can define a map from Z_N^* to itself defined by $f_a(y) = aymodN$. This map is 1-1, since if $ay_1 \equiv_N ay_2$, then $N|a(y_1 - y_2)$, so $N|y_1 - y_2$. Since the only way a map from a finite set to itself can be 1-1 is for it to also be onto, there is a y with $aymodN = 1$. This y is the inverse. We'll see an explicit formula for the inverse later.

A sub-group H of a group G is a subset of G such that if $h_1, h_2 \in H$ then $h_1 * h_2 \in H$, $h_1^{-1} \in H$ and $1 \in H$. (Note that groups are closed under $*$ and always contain the identity element.)

Theorem If G is a finite group and H is a subgroup then $|H|$ is a divisor of $|G|$.

Corollary: If $|H| \neq |G|$, then $|H| \leq 1/2|G|$.

Example: $p = 7$, $G = \{1, 2, 3, 4, 5, 6\}$. Let's suppose $\{1, 4\} \subseteq H$, then closure under $*$ implies $2 \in H$. Then we have $1, 2, 4 \in H$, $3, 5, 6 \notin H$.

If H is a sub-group of G, $g \in G$, then

$$gH = \{gh | h \in H\}$$

is called a coset of H.

Lemma 1 —gH— = —H—

Lemma 2 Define $g \sim g'$ iff $g' \in gH$; then $\sim$ is an equivalence relation. (Proof is a homework exercise.)

Proof of theorem from lemmas: From lemma 2, G is partitioned into disjoint cosets of H. From lemma 1, each coset has the same size as H. Therefore, $|G| = |H| *$ the number of distinct cosets.