**CSE 525 Randomized Algorithms & Probabilistic Analysis** **Winter 2008**

# Lecture 3: January 14

*Lecturer: James R. Lee* *Scribe: Alan Ritter, Bao-Nguyen Nguyen*



## 3.1 Primality Testing (continued)

$\mathbb{Z}_n^* = \{a : \gcd(a, n) = 1\}$ refers to the multiplicative group modulo $n$.

### 3.1.1 Fermat Test

Recall from last time, the Fermat test, which makes use of Fermat's Little Theorem to test if $n$ is prime or composite:

choose $a \in \{1 \dots n-1\}$ uniformly at random
**if** $\gcd(a, n) \neq 1$ **then**
    return(composite)
**else if** $a^{n-1} \not\equiv 1 \pmod{n}$ **then**
    return(composite)
**else**
    return(maybe prime)

Note that the Fermat test has one-sided error. It can mistakenly output "prime" when it's input was in fact composite. Also there is a set of numbers, referred to as Carmichael Numbers, for which this test fails regardless of choice of $a$.

**Claim 3.1** *If $n$ is not a Carmichael Number(CN), then* $\mathbf{Pr}[error] \leq \frac{1}{2}$.

**Proof:** Define $S_n = \{a \in \mathbb{Z}_n^* : a^{n-1} \equiv 1 \pmod{n}\}$. $S_n$ contains the set of inputs which cause Fermat's test to fail. Furthermore $S_n$ is a subgroup of $\mathbb{Z}_n^*$, since it contains 1, and it is closed under multiplication ($(ab)^{n-1} = a^{n-1}b^{n-1} \equiv 1 \pmod{n}$). $S_n$ is also a proper subgroup, because $n$ is not a CN, so there must exist some $a \in \mathbb{Z}_n^*$ such that $a \notin S_n$, so $|S_n| < |\mathbb{Z}_n^*|$. By Lagrange's theorem, $|S_n| \mid |\mathbb{Z}_n^*|$, therefore, $|S_n| \leq \frac{1}{2}|\mathbb{Z}_n^*|$ ■

Claim 3.1 suggests that CNs are the only bad inputs to the Fermat test, and because

$$\lim_{b \to \infty} \mathbf{Pr}[\text{random b bit number is CN}] = 0$$

the Fermat test is useful for some applications.

### 3.1.2 Miller-Rabin Primality test

Miller-Rabin is a randomized primality testing algorithm which does not always fail when it's input is CN.

**Definition 3.2** *$a \in \mathbb{Z}_n^*$ is a quadratic residue (mod n) if $a \equiv x^2$ (mod n) for some $x \in \mathbb{Z}_n^*$. We call $x$ a "square root" of $a$.*

**Claim 3.3** *If $n$ is prime, then $\pm 1$ are the only square roots of 1 (mod n).*

**Proof:** The proof proceeds by contradiction. Assume $x \neq \pm 1$, but $x^2 \equiv 1 \pmod{n}$. Then we have:

$$\begin{aligned} x^2 &\equiv 1 (\text{mod } n) \\ x^2 - 1 &\equiv 0 (\text{mod } n) \\ (x-1)(x+1) &\equiv 0 (\text{mod } n) \end{aligned}$$

Therefore, either $n|x-1$ or $n|x+1$, and thus $x \equiv \pm 1 \pmod{n}$ which contradicts the above assumption. ■

Claim 3.3 suggests another test for primality, namely checking for square roots of 1 (mod $n$) which are not $\pm 1$. The Miller-Rabin algorithm makes use of this in addition to the Fermat test:

```
choose a ∈ {2 ... n − 1} uniformly at random
if gcd(a, n) ≠ 1 then
    return(composite)
choose r, R such that 2^r R = n − 1 (with R odd)
compute b_i = a^{2^i R} for 0 ≤ i ≤ r
if a^{n−1} ≢ 1 (mod n) then
    return(composite)
else if b_0 ≡ 1 (mod n) then
    return(maybe prime)
else
    select the smallest i s.t. b_i = 1
    if b_{i−1} ≢ −1 (mod n) then
        return(composite)
    else
        return(maybe prime)
```

**Claim 3.4** *If $n$ is odd, composite and not a prime power, then $\mathbf{Pr}[error] \leq \frac{1}{2}$.*

It is easy to check if $n$ is a prime power (**the proof is an exercise**).

**Definition 3.5** *$s$ is a bad power if $\exists x(x^s \equiv -1 (\text{mod } n))$. Given a bad power $s$, define $S_n$ as the subgroup of all $x$ s.t. $x^s = \pm 1 (\text{mod } n)$.*

**Lemma 3.6** *If $n$ is composite, odd, and not a prime power, then $S_n$ is a proper subgroup of $\mathbb{Z}_n^*$.*

We first present a proof of claim 3.4 using lemma 3.6:
**Proof:** Let $s^* = 2^{i^*} R$ be the largest bad power. There exists at least one of these, since $R$ is odd, and thus $(-1)^R \equiv -1 \pmod{n}$. Now suppose that $a \in \{2, 3, \ldots, n-1\}$ is not a witness; there are two possible cases:

1. $a^R \equiv a^{2R} \equiv a^{4R} \equiv \ldots \equiv a^{n-1} \equiv 1 \pmod{n}$.
2. For some $i \in \{0, \ldots, r-1\}$, $a^{2^i R} \equiv -1 \pmod{n}$ and $a^{2^{i+1}R} \equiv \ldots \equiv a^{n-1} \equiv 1 \pmod{n}$.

In case 1 above, $a^{s^*} \equiv 1 \pmod{n}$, so $a \in S_n$. In case 2, $2^i R$ is a bad power, and because $s^*$ is the largest bad power, $a^{s^*} = \pm 1 \pmod{n}$. So in either case $a \in S_n$, and because $S_n$ is a proper subgroup of $\mathbb{Z}_n^*$ (due to Lemma 3.6) we can apply Lagrange's Theorem:

$$\begin{aligned} \mathbf{Pr}[\text{error}] = \mathbf{Pr}[\text{a is not a witness}] &\leq \mathbf{Pr}[a \in S_n] \\ &\leq \frac{|S_n|}{|\mathbb{Z}_n^*|} \\ &\leq \frac{1}{2} \end{aligned}$$

■

Now we prove lemma 3.6:
**Proof:** Because $n$ is odd, composite, and not a prime power, there exists $n_1, n_2$ which are co-prime and odd, such that $n = n_1 n_2$. Since $s$ is a bad power, there exists an $x$ s.t. $x^s \equiv -1 \pmod{n}$. Now using the Chinese Remainder Theorem, we can find $y \in \mathbb{Z}_n^*$ such that:

$$\begin{aligned} y &= x \pmod{n_1} \\ y &= 1 \pmod{n_2} \end{aligned}$$

So we have that:

$$y^s = x^s = -1 \pmod{n_1} \quad (3.1)$$
$$y^s = 1 \pmod{n_2} \quad (3.2)$$

If $y$ is in $S_n$, then $y^s = \pm 1 \pmod{n}$. If $y^s = 1 \pmod{n}$, then $y^s = 1 \pmod{n_1}$ which contradicts 3.1. If $y^s = -1 \pmod{n}$, then $y^s = -1 \pmod{n_2}$ which contradicts 3.2. Therefore $y \notin S_n$ by contradiction, so $S_n$ is a proper subgroup of $\mathbb{Z}_n^*$. ■

## 3.2 The Probabilistic Method

The probabilistic method is a nonconstructive, powerful mathematical tool pioneered by Paul Erdos, for proving the existence of a prescribed kind of object. It works by showing that given some probability distributions over random objects, the probability that the object we choose satisfies the desired properties is more than 0.

### 3.2.1 MAX-3SAT

#### 3.2.1.1 Existence of good solution - probabilistic method

MAX-3SAT is an NP-hard optimization problem. This example will show how a simple probabilistic method can yields a good lower bound for this problem.

**Input:** A 3-CNF boolean formula $\varphi = C_1 \wedge C_2 \wedge \cdots \wedge C_m$ in which each $C_i$ is a disjunction of 3 literals on the set of variables $\{x_1, x_2, \cdots x_n\}$. For the purpose of this this section, we assume that the literals of each $C_i$ come from different variables.

**Output:** Find an truth assignment to boolean variables $\{x_1, x_2, \cdots, x_n\}$ to maximize the number of clauses $C_i$ that are satisfied.

**Claim 3.7** *For every $\varphi$, there exists and assignment that satisfies at least $\frac{7m}{8}$ clauses.*

**Proof:** Choose a truth assignment to $\{x_1, x_2, \cdots, x_n\}$uniformly independently at random. Define

$$Y_i = \begin{cases} 1 & \text{if } C_i \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases}$$

Let $Y$ be the number of satisfied clauses: $Y = \sum_{i=1}^{m} Y_i$.

Since among the 8 truth assignments to the variables of $C_i$, exactly one of them doesn't satisfy $C_i$ (recall the assumption that the three literals of $C_i$ come from different variables), we have:

$$\mathbf{E}[Y_i] = \mathbf{Pr}[C_i \text{ is satisfied}] = \frac{7}{8}$$

By linear of expectation,

$$\mathbf{E}[Y] = \sum_{i=1}^{m} \mathbf{E}[Y_i] = \frac{7}{8}m.$$

In the sample space, there exist a point at which $Y$ takes value at least $\mathbf{E}[Y]$. Thus, there exists an assignment satisfying at least $\frac{7}{8}m$ clauses. ■

#### 3.2.1.2 Constructing good solutions - Markov's inequality

The previous section proved that there exists an assignment satisfying at least $\frac{7}{8}m$ clauses without telling how to construct such a solution. In fact, the probability of hitting one may be vanishingly small. In this section, we show that the randomized algorithm which assigns random values to variables produces good solutions - those satisfy at least $\frac{3m}{4}$ clauses - with high probability. To bound the chance of getting good solutions, we need a new tool: Markov's Inequality.

**Theorem 3.8 (Markov's Inequality)** *If $X$ is a non-negative random variable then for all $\alpha > 0$,*

$$\mathbf{Pr}[X \geq \alpha \mathbf{E}[X]] \leq \frac{1}{\alpha}$$

**Proof:** The proof is simple and is left as an exercise. ■

Let $Z = m - Y$, the number of unsatisfied clauses. Then $Z$ is non-negative and

$$\mathbf{E}[Z] = m - \mathbf{E}[Y] = \frac{m}{8}$$

By Markov's Inequality,

$$\mathbf{Pr}[Z \geq \alpha \mathbf{E}[Z]] = \mathbf{Pr}[Z \geq \frac{\alpha m}{8}] \leq \frac{1}{\alpha}$$

Choose $\alpha = 2$, we have

$$\mathbf{Pr}[Z \geq \frac{m}{4}] \leq \frac{1}{2}$$

or in other words,

$$\mathbf{Pr}[Y \geq \frac{3m}{4}] \geq \frac{1}{2}$$

Therefore, a random assignment satisfies at least $\frac{3}{4}$ of the clauses with the probability at least $\frac{1}{2}$.