

# ALGORITHMS FOR NUMBER-THEORY PROBLEMS

We will cover:

- Divisibility
  - Euclid's algorithm for the greatest common divisor (gcd)
- Congruent structures (modular arithmetic)
  - $\mathbb{Z}_n$ ,  $\mathbb{Z}_n^*$ , inverse computation
  - The Chinese remainder theorem
- APPLICATIONS
  - RSA Cryptosystem
  - Miller & Rabin's randomized PRIMALITY testing

## (INTEGER) DIVISIBILITY

DEF Let  $d \in \mathbb{Z} - \{0\}$ ,  $a \in \mathbb{Z}$ . We write  $d|a$   
(read  $d$  divides  $a$ ) if  $\exists k \in \mathbb{Z} : a = k \cdot d$   
( $a$  multiple of  $d$ )

FACT  $d|a \Rightarrow -d|a$  ( $k \rightarrow -k$ )

DEF If  $d|a$  and  $d > 0$ ,  $d$  is called a  
divisor of  $a$

## PROPERTIES (prove as an exercise)

a.  $\forall d \in \mathbb{Z} - \{0\} : d|0$  ( $k=0$ )

b.  $(a \neq 0) \wedge (d|a) \Rightarrow |d| \leq |a|$

c.  $(a|b) \wedge (b|c) \Rightarrow a|c$  (or  $|ab| = |ac|$ )

DEF A positive integer  $p > 1$  is prime if:  
 $\forall d > 0 : (d|p) \Rightarrow (d=1) \vee (d=p)$   
( $p$  admits only 1 and  $p$  as divisors)

### DIVISION THEOREM (no proof :-)

$\forall a \in \mathbb{Z}, \forall n \in \mathbb{Z}^+ :$   
 $\exists ! q, r : a = q \cdot n + r$ , with  
 uniqueness  $q, r \in \mathbb{Z}$  and  $0 \leq r < n$

$q$  is called quotient (of the integer division of  $a$  by  $n$ )

$r$  is called remainder (of the integer division of  $a$  by  $n$ )

NOTATION :  $q = \lfloor a/n \rfloor, r = a \bmod n$

e.g.)  $\frac{-1}{4} = (-1) \cdot n + (n-1)$   
 $\uparrow$   
 $a \Rightarrow -1 \bmod 4 = 3$   
 $\lfloor -1/4 \rfloor = -1$

### DEF

$d > 0$  is common divisor of  $a$  and  $b$ , with  $a, b \in \mathbb{Z}$ , if  $(d|a) \wedge (d|b)$

## PROPERTIES OF COMMON DIVISORS (proof: exercise)

- $(d|a) \wedge (d|b) \Rightarrow d| (a+b)$
- More generally,  $\forall x, y \in \mathbb{Z}$ :  
 $(d|a) \wedge (d|b) \Rightarrow d| (ax+by)$

**DEF** Let  $a, b \in \mathbb{Z}$ , with  $|a| + |b| > 0$  (at least one  $\neq 0$ ) - We let  
 $\gcd(a, b) = \max \{d > 0 : (d|a) \wedge (d|b)\}$

**NOTE** Since either  $a$  or  $b$  is  $> 0$ , the set  $\{d > 0 : (d|a) \wedge (d|b)\}$  is finite, since  $d \leq |a|$  if  $a \neq 0$  or  $d \leq |b|$  if  $b \neq 0$

**CONVENTION**  $\Rightarrow |\{d > 0 : d|ab\}| \rightarrow \infty$

We set  $\gcd(0, 0) = 0$

## PROPERTIES OF gcd (prove as exercise)

- H 1. If  $a, b \neq 0$ :  $1 \leq \gcd(a, b) \leq \min \{|a|, |b|\}$
- O 2.  $\gcd(a, 0) = |a|$  symmetric, sign insensitive
- M 3.  $\gcd(a, b) = \gcd(b, a) = \gcd(-a, b) = \gcd(|a|, |b|)$
- E 4.  $\gcd(a, ka) = |a| \quad \forall a, k \in \mathbb{Z}$

Important characterization of  $\gcd(a, b)$

**THEOREM (Bezout identity) (1779)**

Let  $|a| + |b| > 0$ . Then

$$\gcd(a, b) = \min \{d > 0 \mid \exists x, y \in \mathbb{Z} \mid d = ax + by\}$$

In words:  $\gcd(a, b)$  is the minimum positive linear combination of  $a$  and  $b$  with integer coefficients

Important characterization of  $\gcd$  with many applications

**PROOF** Let  $S = \min \{ d > 0 \mid \exists x, y \in \mathbb{Z} \mid d = ax + by \}$   
 We will prove  $\gcd(a, b) \leq S$

$\leq$  By the property of common divisors we know that  $\gcd(a, b) \mid ax + by \quad \forall x, y \in \mathbb{Z}$   
 Since  $S = \min \{ d > 0 \mid \exists x, y \in \mathbb{Z} \mid d = ax + by \} = 2\bar{x} + b\bar{y}$   
 we have  $\gcd(a, b) \mid S$ . Also  $S > 0$ , thus  $0 < \gcd(a, b) \leq |S| = S$

$\geq$  We'll prove that  $S \mid ax + by \quad \forall x, y \in \mathbb{Z}$  &

If this is the case:  $S \mid a$  ( $x=1, y=0$ ) and  $S \mid b$  ( $x=0, y=1$ ) - Thus  $S$  is a common divisor of  $a$  and  $b$ , therefore  $\gcd(a, b) \geq S$ .

$\rightarrow$  Say that  $S = 2\bar{x} + b\bar{y}$ . Let  $C = ax + by$ , for arbitrary  $x, y \in \mathbb{Z}$ .

Apply the division theorem to  $(C, S)$ :

$$\exists q, r : C = ax + by = q \cdot S + r \quad 0 \leq r < S$$

$$\begin{aligned} \text{We have that } r &= C - qS = ax + by - q(2\bar{x} + b\bar{y}) \\ &= 2(x - q\bar{x}) + b(y - q\bar{y}) = 2x'' + b y'' \end{aligned}$$

Then  $r=0$ , or otherwise we would have found  $0 < r = 2x'' + by'' < s$ , which contradicts  $s = \min\{dsol\exists x, y \in \mathbb{Z} | dx+by\}$ .

Since  $r=0$  we have that  $s \leq ax+by$ !

Efficient computation of  $\gcd(a, b)$

From now on, we restrict to the case  $a, b \geq 0$  ( $\gcd(a, b) = \gcd(|a|, |b|)$ )

Euclid's algorithm (300 BC) :

```
EUCLID (a, b)
③ if (b=0) then return a
      return EUCLID(b, a mod b)
```

Simple recursive algorithm.

CORRECTNESS

①  $\text{EUCLID}(a, b)$  terminates.

In each call the second parameter decreases (from  $b$  to  $a \text{ mod } b < b$ )  $\Rightarrow$  No more than  $b+1$  calls.

NOTE. If  $a < b$ ,  $\text{EUCLID}(a, b)$  calls  $\text{EUCLID}(b, a)$ . In subsequent calls the first parameter is always  $\geq$  than the second.

②  $\text{EUCAL}(a, b)$  returns  $\text{gcd}(a, b)$   
 This is true for  $b=0$ , since  
 $\text{gcd}(a, 0) = a$ . It remains to  
 prove that when  $b > 0$ ,

$$\text{gcd}(a, b) = \text{gcd}(b, a \bmod b)$$

≤  
≥

≤ by Eratosthenes' identity:

- ①  $\text{gcd}(a, b) = \min\{d > 0 \mid \exists x, y \in \mathbb{Z} : d = ax + by\}$
- ②  $\exists x', y' : \text{gcd}(b, a \bmod b) = b x' + (a \bmod b) y'$

Let  $a = qb + r \bmod b$  (by division theorem)  
 Then

$$a \bmod b = a - qb, \text{ hence from ② :}$$

$$\begin{aligned} \exists x', y' : 0 < \text{gcd}(b, a \bmod b) &= b x' + (a - qb) y' \\ &= a y' + (x' - q y') b \end{aligned}$$

We have written  $\text{gcd}(b, a \bmod b)$  as  
 a positive integer linear combination  
 of  $a$  and  $b$ . Therefore, from ①

$$\text{gcd}(a, b) \leq \text{gcd}(b, a \bmod b)$$

// Let  $d' = \gcd(b, a \bmod b)$ . We

have  $d' | b$  and  $d' | a \bmod b$

hence  $d' | x \cdot b + y \cdot a \bmod b \quad \forall x, y \in \mathbb{Z}$

Recall that  $a = qb + r \bmod b$

and set  $x = q$  and  $y = 1$ !

$$d' | qb + r \bmod b = r$$

But then

$$(d' | b) \wedge (d' | r) \Rightarrow$$

$$\gcd(a, b) \geq d' = \gcd(b, a \bmod b)$$

## RUNNING TIME

Let us evaluate the number of recursive calls performed by  $\text{EUCLID}(a, b)$  in the worst case.

We can assume that  $a > b > 0$ .

otherwise:

1.  $b=0 \Rightarrow$  one call (base case)

2.  $a=b \Rightarrow$  two calls:

$$\text{EUCLID}(a, b) \rightarrow \text{EUCLID}(b, b) \rightarrow \text{EUCLID}(b, 0)$$

3.  $a < b \Rightarrow \text{EUCLID}(a, b) \rightarrow \text{EUCLID}(b, a)$

and the condition is established by swapping  $a$  and  $b$  (+1 extra call)

The analysis relies on Fibonacci numbers

$$F_1 = F_2 = 1, \quad F_k = F_{k-1} + F_{k-2} \quad k > 2$$
$$1, 1, 2, 3, 5, 8 \dots$$

It holds :

$$\begin{aligned} F_k &= \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^k - \frac{1}{\sqrt{5}} \left( \frac{1-\sqrt{5}}{2} \right)^k \\ &\geq \frac{1}{2\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^k > 0 \end{aligned}$$

(exponential in  $k$  ( $\frac{1+\sqrt{5}}{2} \approx 1.62$ ))

We prove that for EUCLID to perform many recursive calls, the inputs must be very large. (thus, the number of calls is "small" w.r.t. the size of the inputs)

LEMMA For  $a > b > 0$ , if EUCLID( $a, b$ ) makes  $K$  calls overall (including the external one), then

$$a \geq F_{K+2} \text{ and } b \geq F_{K+1}$$

We will get a bound on the number of calls for a specific input size by inverting the inequality (solving in  $K$ ).

(later)

PROOF By induction on K.

BASE K=1: Since  $b > 0 \Rightarrow b \geq 1 = F_2$

Also,  $a > b \Rightarrow a > 1 \Rightarrow a \geq 2 = F_3$

HP True for K-1 cells, K>1

TH If EUCLID(a,b) makes K cells,  
then EUCLID(b, a mod b) (first re -  
ursive call) makes K-1 cells.

By HP:  $b \geq F_{(K-1)+2} = F_{K+1}$ , (TH for b)

$$a \text{ mod } b \geq F_{(K-1)+1} = F_K$$

However  $d = q \cdot b + r \text{ mod } b$ , with  
 $q \geq 1 (a > b)$ . Thus:

$$a \geq b + r \geq F_{K+1} + F_K = F_{K+2} \quad (\text{TH for } a)$$

Let  $\bar{K}$ :  $F_{\bar{K}} < b \leq F_{\bar{K}+1}$ . Then F is  
an upper bound to the number of  
cells performed by EUCLID(a,b)

We have  $b > F_{\bar{K}} \geq \frac{1}{2\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^{\bar{K}} \Rightarrow$

$$\log_{\frac{1+\sqrt{5}}{2}} b > \bar{K} + c \quad (c = \log_{\frac{1+\sqrt{5}}{2}} \frac{1}{2\sqrt{5}} < 0)$$

$$\Rightarrow \bar{K} < \log_{\frac{1+\sqrt{5}}{2}} b - c \Rightarrow \bar{K} = O(\log b)$$

Since

$$T = O(\log b) = O(|\langle a, b \rangle|)$$

EUCLID( $a, b$ ) makes at most a LINEAR number of calls in the size  $|\langle a, b \rangle|$ .

Each call must compute  $a \bmod b$

which takes time  $O(|\langle a, b \rangle|^2)$  (exercise: implement integer division) (in class)

$$\Rightarrow T_{\text{EUCLID}}(|\langle a, b \rangle|) = O(|\langle a, b \rangle|^3) !$$

## EXTENDED EUCLID ALGORITHM

EUCLID( $a, b$ ) can be easily extended to compute the integer coefficients of Bezout's identity:

$$\gcd(a, b) = ax' + by', x', y' \in \mathbb{Z}$$

These coefficients have important applications (e.g., RSA).

Extension:

$$\text{BASE CASE } b=0 \Rightarrow \gcd(a, b)=a$$

$$\Rightarrow x'=1, y'=0 \quad (a = a \cdot 1 + b \cdot 0)$$

REMARK: any other choice for  $\bar{y}'$  would be fine: coefficients are not unique

SUBSTRUCTURE: Suppose that we are given:

$$\gcd(b, a \bmod b) = b\bar{x} + (a \bmod b)\bar{y}$$

Since  $\gcd(a, b) = \gcd(b, a \bmod b)$  we rewrite  $b\bar{x} + (a \bmod b)\bar{y}$  as a linear combination of  $a$  and  $b$ .

We know that  $a = \lfloor a/b \rfloor \cdot b + a \bmod b$

(Division theorem)

$$a \bmod b = (a - \lfloor a/b \rfloor b)$$

Substituting:

$$\begin{aligned} \gcd(a, b) (\equiv \gcd(b, a \bmod b)) &= \\ &= b\bar{x} + (a - \lfloor a/b \rfloor b)\bar{y} = \\ &= a\bar{y} + b(\bar{x} - \lfloor a/b \rfloor \bar{y}) \\ &\quad \uparrow \qquad \uparrow \\ &\quad x' \qquad y' \end{aligned}$$

The algorithm immediately follows  
**EXTENDED-EUCLID** ( $a, b$ ) (EE( $a, b$ ) for short)  
 returns the struct  $\{\gcd(a, b), (x', y')\}$

$\text{EE}(a, b)$   
 $\beta$  if  $(b=0)$   
 then return  $\{0, (1, 0)\}$   
 $R \{d, (\bar{x}, \bar{y})\} \leftarrow \text{EE}(b, a \bmod b)$   
 $C$  return  $\{d, (\bar{y}, \bar{x} - \lfloor a/b \rfloor \bar{y})\}$

**EXAMPLE**     $\text{EE}(24, 16) \rightarrow \{8, (1, -1 \cdot 1) = (1, -1)\}$   
 $\downarrow$   
 $\text{EE}(16, 8) \rightarrow \{8, (0, 1 - 2 \cdot 0) = (0, 1)\}$   
 $\downarrow$   
 $\text{EE}(8, 0) \rightarrow \{8, (1, 0)\}$

Indeed  $\gcd(24, 16) = 8 = 24 \cdot 1 + 16 \cdot (-1)$

