CSCI 2200
FOUNDATIONS OF COMPUTER SCIENCE

David Goldschmidt
goldsd3@rpi.edu
Fall 2022

PI MU EPSILON – MATH HONOR SOCIETY TALK

Presentation about proofs tomorrow (Wednesday, October 26) at 7:00PM…

…in Mothers (bottom floor of the Rensselaer Union)

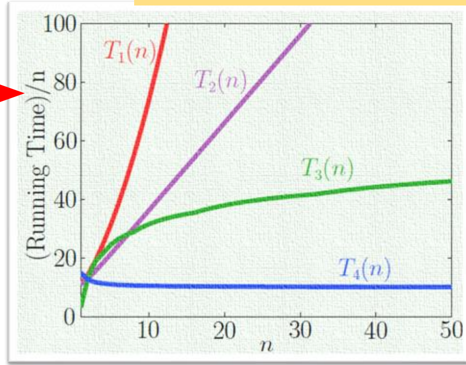
Relevant and interesting to students taking CSCI 2200, MATH 4090, etc.

You should attend — celebrate your work now that we're halfway through the semester!

CATEGORIZING ALGORITHM RUNTIMES

Which algorithm is best…?

We want to know how each algorithm scales with input size *n*, so divide each runtime formula by *n*…



$$T_1(n) = 2 + \tfrac{31}{6}n + \tfrac{7}{2}n^2 + \tfrac{1}{3}n^3$$
$$T_2(n) = 2 + 6n + 3n^2$$
$$3n(\log_2 n + 1) - 9 \leq T_3(n) \leq 12n(\log_2 n + 3) - 9$$
$$T_4(n) = 5 + 10n$$

What if we further improve $T_4(n) = 50 + 8n$…?

CATEGORIZING ALGORITHM RUNTIMES

```
 lim      T (n)            lim      5 + 10n
           4
n->inf   ------    =     n->inf    ---------
           n                          n

                lim       5       10n
        =                ---  +  -----   =   10
              n->inf      n        n
```

(in-lecture notes)

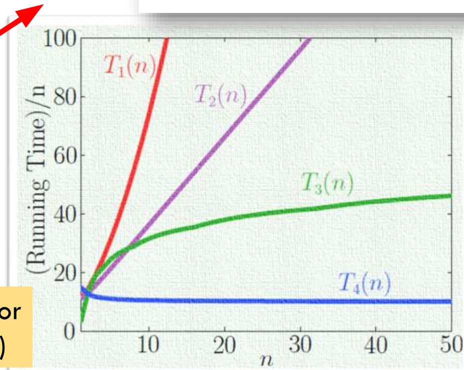
We focus on *scaling up*, i.e., when input size *n* grows very large — when $n \to \infty$

Algorithm 4 is *linear* in *n*…

…as $n \to \infty$, $\dfrac{T_4(n)}{n} \to c$ (a constant)

Therefore, we state that $T_4(n) \in \Theta(n)$

Try doing this for $T_1(n)$ and $T_2(n)$



$$T_1(n) = 2 + \tfrac{31}{6}n + \tfrac{7}{2}n^2 + \tfrac{1}{3}n^3$$
$$T_2(n) = 2 + 6n + 3n^2$$
$$3n(\log_2 n + 1) - 9 \leq T_3(n) \leq 12n(\log_2 n + 3) - 9$$
$$T_4(n) = 5 + 10n$$

We categorize algorithms based on *growth rates* of their runtimes, which makes it easier for us to describe them in comparison with one another

```
         T (n)
lim       4             lim      5 + 10n
       ------   =              ---------
n->inf    n            n->inf      n


         lim      5      10n
   =            --- + -----   =  10 (a constant)
        n->inf    n       n


         T (n)                              2
lim       2             lim      2 + 6n + 3n
       ------   =              ---------------
n->inf    n            n->inf         n

         lim      2
   =            --- + 6 + 3n  =  infinity
        n->inf    n
```

```
T (n) is NOT in Big-Theta(n), i.e., not linear
 2

T (n) is in little-omega(n), i.e., T (n) > linear
 2                                   2


         T (n)                              2
lim       2             lim      2 + 6n + 3n
       ------   =              ---------------
n->inf    2            n->inf          2
         n                            n

         lim       2       6
   =             ---- + --- + 3  =  3 (a constant)
        n->inf     2       n
                  n
                                2
T (n) is in Big-Theta(n ), i.e., quadratic
 2
```

ASYMPTOTICALLY LINEAR FUNCTIONS – Θ(n)

Recurrence $T \in \Theta(n)$ if there are positive constants c and C such that…

$$c \cdot n \leq T(n) \leq C \cdot n$$

As n grows toward ∞, dividing T by n will give us 0, a constant, or ∞

$$\frac{T(n)}{n} \xrightarrow[n\to\infty]{} \begin{cases} \infty & T \in \omega(n), \quad \text{“}T > n\text{”}; \\ \text{constant}>0 & T \in \Theta(n), \quad \text{“}T = n\text{”}; \\ 0 & T \in o(n), \quad \text{“}T < n\text{”}. \end{cases}$$

little-omega-of-n

big-theta-of-n

little-oh-of-n

ASYMPTOTICALLY LINEAR FUNCTIONS – Θ(n)

Example functions that are *asymptotically* linear, i.e., that are in $\Theta(n)$…

$2n + 7$ $30n + 10^{100}$ $2n + 15\sqrt{n}$ $10^9 n + 3$ $2n + \log n$

Functions that are not asymptotically linear, i.e., that are not in $\Theta(n)$…

$10^{-9}n^2$ $n^{1.0001}$ $n^{0.9999}$ $n \log n$ 2^n

How do we know if $T(n) \in \Theta(n)$…?

ASYMPTOTICALLY LINEAR FUNCTIONS – Θ(n)

Example functions that are *asymptotically* linear, i.e., that are in $\Theta(n)$…

$2n + 7$ $30n + 10^{100}$ $2n + 15\sqrt{n}$ $10^9 n + 3$ $2n + \log n$

Functions that are not asymptotically linear, i.e., that are not in $\Theta(n)$…

$10^{-9}n^2$ $n^{1.0001}$ $n^{0.9999}$ $n \log n$ 2^n

also see if you can determine constants c and C for those that are asymptotically linear

How do we know if $T(n) \in \Theta(n)$…? Divide by n and then take the limit to ∞

We can generalize this to any function $f(n)$ — not just the linear $f(n) = n$

$$2n + 15n^{1/2}$$

$$\mathrm{sqrt}(n) = n^{1/2}$$

$$\lim_{n \to \inf} \frac{2n + 15n^{1/2}}{n} = \lim_{n \to \inf} 2 + \frac{15}{n^{1/2}}$$

$$= 2 \text{ (a constant)}$$

therefore, $2n + 15n^{1/2}$ is in Big-Theta(n) or is asymptotically linear

what are upper and lower bounds c and C here...?

$$c\ n \;\; <= \;\; T(n) \;\; <= \;\; C\ n$$

set c to 2 and C to 17, we have

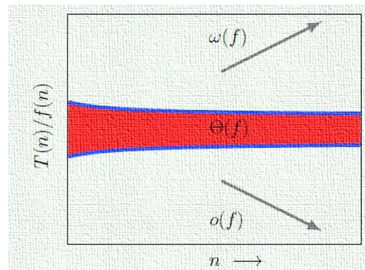
$$2\ n \;\; <= \;\; T(n) \;\; <= \;\; 17\ n$$

GENERAL ASYMPTOTIC FUNCTIONS

We can generalize this to any function $f(n)$ — not just the linear $f(n) = n$

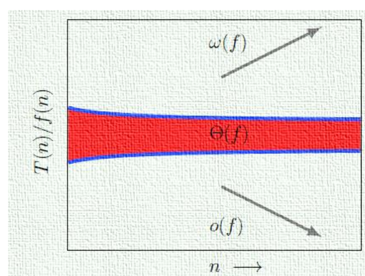
As n grows toward ∞, dividing T by $f(n)$ will again give us 0, a constant, or ∞

$$\frac{T(n)}{f(n)} \xrightarrow[n\to\infty]{} \begin{cases} \infty & T \in \omega(f),\ \text{“}T > f\text{”}; \\ \text{constant}>0 & T \in \Theta(f),\ \text{“}T = f\text{”}; \\ 0 & T \in o(f),\ \text{“}T < f\text{”}. \end{cases}$$



GENERAL ASYMPTOTIC FUNCTIONS

$$\frac{T(n)}{f(n)} \xrightarrow[n\to\infty]{} \begin{cases} \infty & T \in \omega(f),\ \text{“}T > f\text{”}; \\ \text{constant}>0 & T \in \Theta(f),\ \text{“}T = f\text{”}; \\ 0 & T \in o(f),\ \text{“}T < f\text{”}. \end{cases}$$



$T \in o(f)$	$T \in O(f)$	$T \in \Theta(f)$	$T \in \Omega(f)$	$T \in \omega(f)$
“$T < f$”	“$T \le f$”	“$T = f$”	“$T \ge f$”	“$T > f$”
	$T(n) \le Cf(n)$	$cf(n) \le T(n) \le Cf(n)$	$cf(n) \le T(n)$	

FREQUENTLY OCCURRING GROWTH RATES

Runtimes that are reasonable…

log	linear	loglinear	quadratic	cubic
$\Theta(\log n)$	$\Theta(n)$	$\Theta(n \log n)$	$\Theta(n^2)$	$\Theta(n^3)$

best → worst

Runtimes that are unreasonable…

superpolynomial	exponential	factorial	forget it…
$\Theta(n^{\log n})$	$\Theta(2^n)$	$\Theta(n!)$	$\Theta(n^n)$

→ worst

TRICKS TO DETERMINING GROWTH RATE

For polynomials, focus on the highest order term to determine the growth rate…

$2n^2$	$n^2 + n\sqrt{n}$	$n^2 + \log^{256} n$	$n^2 + n^{1.99} \log^{256} n$
$\Theta(n^2)$	$\Theta(n^2)$	$\Theta(n^2)$	$\Theta(n^2)$

Divide by n^2 and take the limit to ∞ to verify…

For summations, the growth rate is the number of *nestings* plus the order of the *summand*…

$\sum_{i=1}^{n} i$	$\sum_{i=1}^{n} \sum_{j=1}^{i} 1$	$\sum_{i=1}^{n} \sum_{j=1}^{i} ij$
$\Theta(n^2)$	$\Theta(n^2)$	$\Theta(n^4)$

Remove the summations by determining an equivalent $f(n)$, then divide by n^2 and take the limit to ∞ to verify…

(or n^4)

NUMBER THEORY – DIVISIBILITY

Given $n = 27$ and $d = 7$, what is the minimum non-negative remainder r such that…

$n = qd + r$ for quotient $q \in \mathbb{Z}$?

Tinker with $q = -1, 0, 1, 2, 3, \ldots$

Here, $q = 3$ and the remainder is $r = 6$ — i.e., $r = \text{rem}(n, d) = \text{rem}(27, 7) = 6$

Quotient-Remainder Theorem
For $n \in \mathbb{Z}$ and $d \in \mathbb{N}$, $n = qd + r$. The quotient $q \in \mathbb{Z}$ and remainder $0 \leq r < d$ are *unique*.

We define *divisibility* by stating d divides n (or $d \mid n$) iff $n = qd$ for some $q \in \mathbb{Z}$…

…in other words, remainder $\text{rem}(n, d) = 0$

e.g., $3 \mid 12$ or $7 \mid 42$ or $101 \mid 808$ or $x \mid xy$ or …

NUMBER THEORY – DIVISIBILITY

We define *divisibility* by stating d divides n (or $d \mid n$) iff $n = qd$ for some $q \in \mathbb{Z}$

$$d \mid n \leftrightarrow n = dk \text{ for } k \in \mathbb{Z}$$

Use the above equivalence to prove the claims below…

(a) $d \mid 0$

(b) if $d \mid m$ and $d' \mid n$, then $dd' \mid mn$

(c) if $d \mid m$ and $m \mid n$, then $d \mid n$

(d) if $d \mid n$ and $d \mid m$, then $d \mid (m + n)$

(e) if $d \mid n$, then $xd \mid xn$ for $x \in \mathbb{N}$

(f) if $d \mid (m + n)$ and $d \mid m$, then $d \mid n$

NUMBER THEORY – PRIME NUMBERS

We can define the set of *prime* numbers P using divisibility…

note that 1 is not prime

$P = \{\, p \mid p \geq 2$ with positive divisors 1 and p, i.e., $x \mid p$ iff $x = 1$ or $x = p \,\}$

The *Fundamental Theorem of Arithmetic* states that for all natural numbers $n \geq 2$, we can write n as the product of one or more prime numbers

We proved this theorem using strong induction…

…but we did not prove the *uniqueness* of these products (aside from reordering)

e.g., $43 \times 47 = 2021$ is unique; $7 \times 17 \times 17 = 2023$ is unique; etc.

No other group of prime numbers will produce 2021 or 2023 when multiplied together!

GREATEST COMMON DIVISOR (GCD)

i.e., $d \mid m \wedge d \mid n \rightarrow d \leq \gcd(m, n)$

Definition. Greatest Common Divisor, GCD

Let m, n be two integers not both zero. $\gcd(m, n)$ is the largest integer that divides both m and n: $\gcd(m,n) \mid m$, $\gcd(m,n) \mid n$ and any other common divisor $d \leq \gcd(m, n)$.

What is gcd(30, 42)?

Divisors of 30 are {1, 2, 3, 5, 6, 15, 30}

Divisors of 42 are {1, 2, 3, 6, 7, 14, 21, 42}

Common divisors: {1, 2, 3, 6}

Therefore, gcd(30, 42) = 6

Note that $\gcd(m, n) = \gcd(n, m)$…

What is gcd(30, 49)?

Divisors of 30 are {1, 2, 3, 5, 6, 15, 30}

Divisors of 49 are {1, 7}

Common divisors: {1}

Therefore, gcd(30, 49) = 1

Relatively Prime

If $\gcd(m, n) = 1$, then m, n are relatively prime.

Definition. Greatest Common Divisor, GCD

Let m, n be two integers not both zero. $\gcd(m, n)$ is the largest integer that divides both m and n: $\gcd(m, n) | m$, $\gcd(m, n) | n$ and $d | m \wedge d | n \rightarrow d \leq \gcd(m, n)$.

GREATEST COMMON DIVISOR (GCD)

Theorem.

$\gcd(m, n) = \gcd(\text{rem}(n, m), m)$.

Prove this theorem using a direct proof…

Proof. We prove the claim using a direct proof.

Here, $r = \text{rem}(n, m)$ means $n = qm + r$, so $r = n - qm$.

Let $D = \gcd(m, n)$ and $d = \gcd(m, r)$.

Since $D | m$ and $D | n$, we have $D | (n - qm)$ or $D | r$.

Therefore, $D \leq \gcd(m, r) = d$ — i.e., D is a common divisor of m and r.

For d, we have $d | m$ and $d | r$, so $d | qm + r$ or $d | n$.

Therefore, $d \leq \gcd(m, n) = D$ — i.e., d is a common divisor of m and n.

Since $D \leq d$ and $D \geq d$, it must follow that $D = d$; thus, $\gcd(m, n) = \gcd(m, r)$. ■

EUCLID'S ALGORITHM – RECURSIVE FUNCTION

```
/* invariant: m < n */
gcd( m, n )
{
   if ( m == 0 )
     return n;
   else
     return gcd( n, rem( m, n ) );
}
/********************* m % n ****/
```

EUCLID'S ALGORITHM – EXAMPLE

Theorem.
$\gcd(m, n) = \gcd(\text{rem}(n, m), m)$.

$$\begin{aligned}
\gcd(42, 108) &= \gcd(24, 42) && \Rightarrow 24 = 108 - 2 \times 42 \\
&= \gcd(18, 24) && \Rightarrow 18 = 42 - 24 = 42 - \underbrace{(108 - 2 \times 42)}_{24} = 3 \times 42 - 108 \\
&= \gcd(6, 18) && \Rightarrow 6 = 24 - 18 = \underbrace{(108 - 2 \times 42)}_{24} - \underbrace{(3 \times 42 - 108)}_{18} = 2 \times 108 - 5 \times 42 \\
&= \gcd(0, 6) && \Rightarrow 0 = 18 - 3 \times 6 \\
&= 6 && [\text{since } \gcd(0, n) = n]
\end{aligned}$$

EUCLID'S ALGORITHM – EXAMPLE

Theorem.
$\gcd(m, n) = \gcd(\text{rem}(n, m), m)$.

Each remainder is a *linear combination* of the original *m* and *n* (i.e., of 42 and 108)...

$$\begin{aligned}
\gcd(42, 108) &= \gcd(24, 42) && \Rightarrow 24 = \mathbf{108} - 2 \times \mathbf{42} \\
&= \gcd(18, 24) && \Rightarrow 18 = 42 - 24 = 42 - \underbrace{(108 - 2 \times 42)}_{24} = 3 \times \mathbf{42} - \mathbf{108} \\
&= \gcd(6, 18) && \Rightarrow 6 = 24 - 18 = \underbrace{(108 - 2 \times 42)}_{24} - \underbrace{(3 \times 42 - 108)}_{18} = 2 \times \mathbf{108} - 5 \times \mathbf{42} \\
&= \gcd(0, 6) && \Rightarrow 0 = 18 - 3 \times 6 \\
&= 6 && [\text{since } \gcd(0, n) = n]
\end{aligned}$$

$\gcd(m, n) = mx + ny$ for integers x and y

With $m = 42$ and $n = 108$, can you come up with integers x and y such that $mx + ny < 6$...?

$$\gcd(42, 108) = 6 = \underbrace{2 \times 108 - 5 \times 42}_{\text{linear combination}} = 2n - 5m \text{ with } m = 42 \text{ and } n = 108$$

Using $m = 6$ and $n = 15$, list some positive linear combinations $z = mx + ny$ for integers x and y… (and can you minimize z?)

EUCLID'S ALGORITHM – EXAMPLE

Theorem.
$\gcd(m, n) = \gcd(\text{rem}(n, m), m)$.

Each remainder is a *linear combination* of the original m and n (i.e., of 42 and 108)…

$\gcd(m, n) = mx + ny$ for integers x and y

$$\begin{aligned}
\gcd(42, 108) &= \gcd(24, 42) && \Rightarrow 24 = \mathbf{108} - 2 \times \mathbf{42} \\
&= \gcd(18, 24) && \Rightarrow 18 = 42 - 24 = 42 - \underbrace{(108 - 2 \times 42)}_{24} = 3 \times \mathbf{42} - \mathbf{108} \\
&= \gcd(6, 18) && \Rightarrow 6 = 24 - 18 = \underbrace{(108 - 2 \times 42)}_{24} - \underbrace{(3 \times 42 - 108)}_{18} = 2 \times \mathbf{108} - 5 \times \mathbf{42} \\
&= \gcd(0, 6) && \Rightarrow 0 = 18 - 3 \times 6 \\
&= 6 && [\text{since } \gcd(0, n) = n]
\end{aligned}$$

With $m = 42$ and $n = 108$, can you come up with integers x and y such that $mx + ny < 6$…?

$$\gcd(42, 108) = 6 = \underbrace{2 \times 108 - 5 \times 42}_{\text{linear combination}} = 2m - 5n \text{ with } m = 42 \text{ and } n = 108$$

Using $m = 6$ and $n = 15$, list some positive linear combinations $z = mx + ny$ for integers x and y… (and can you minimize z?)

EUCLID'S ALGORITHM – EXAMPLE

```
Using m = 6 and n = 15, list some positive linear
 combinations z = mx + ny for integers x and y...

  The goal is to minimize z (but keep z > 0)

if x = 1 and y = 1, then z = 6 + 15 = 21

if x = 2 and y = 2, then z = 12 + 30 = 42

if x = -2 and y = 1, then z = mx + ny is the minimum z

                          z =  -12 + 15 = 3

                    gcd( m, n ) = gcd( 6, 15 ) = 3
```

EUCLID'S ALGORITHM TO BEZOUT'S IDENTITY

From Euclid's Algorithm, gcd(m, n) = $mx + ny$ for integers x and y

Let $z = mx + ny > 0$.

Can we find a smaller $z > 0$ that is a linear combination of m and n? No…!

Theorem. Bezout's Identity

$\gcd(m, n)$ is the *smallest positive integer linear combination* of m and n:

$$\gcd(m, n) = mx + ny \qquad \text{for } x, y \in \mathbb{Z}.$$

Bezout's Identity is essentially a "formula" for GCD

GCD FACTS

(i) $\gcd(m, n) = \gcd(m, \text{rem}(n, m))$.

(ii) Every common divisor of m, n divides $\gcd(m, n)$.

(iii) For $k \in \mathbb{N}$, $\gcd(km, kn) = k \cdot \gcd(m, n)$.

(iv) IF $\gcd(l, m) = 1$ AND $\gcd(l, n) = 1$, THEN $\gcd(l, mn) = 1$.

(v) IF $d \mid mn$ AND $\gcd(d, m) = 1$, THEN $d \mid n$.

We can prove (iii)-(v) using Bezout's Identity, e.g., for (iii)…

Proof. We prove the claim that for $k \in \mathbb{N}$, gcd(km, kn) = $k \times$ gcd(m, n).

Here, gcd(km, kn) = $kmx + kny = k(mx + ny)$.

From Bezout's Identity, the RHS must be the smallest possible, i.e., there is no smaller linear combination of m and n.

Use Bezout's Identity to prove claims (iv) and (v)…

Therefore, since $k > 0$, we conclude that gcd(m, n) = $mx + ny$. ■

https://www.youtube.com/watch?v=2vdF6NASMiE

USING BEZOUT'S IDENTITY TO AVOID DISASTER

What does Bezout's Identity have to do with this?

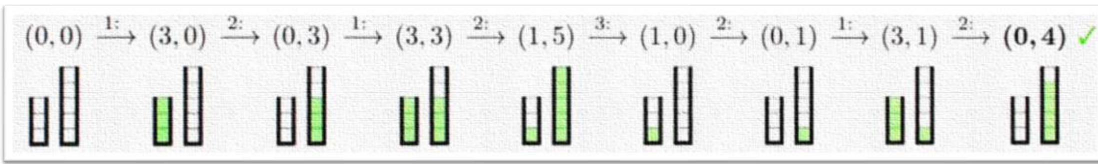
Given only 3- and 5-gallon containers, can we measure exactly 4 gallons?

Strategy comes from the following rules/operations…

1: Repeatedly fill the 3-gallon container (from an unlimited water supply)

2: Pour as much as you can from the 3-gallon container into the 5-gallon container

3: If the 5-gallon container is full, we can empty it



USING BEZOUT'S IDENTITY TO AVOID DISASTER

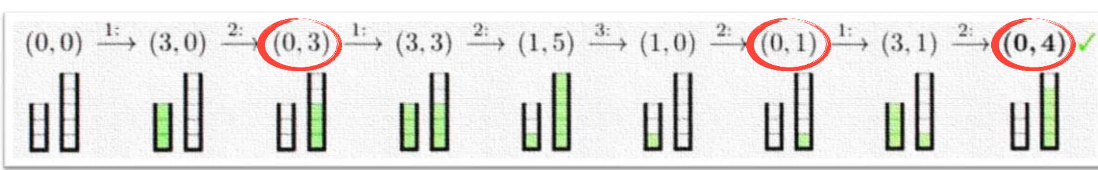
What does Bezout's Identity have to do with this?

When we empty the 3-gallon container into the 5-gallon container,
the state of the problem becomes $(0, \ell)$ with $0 \leq \ell \leq 5$.

Let $x \geq 0$ be the number of times we successfully empty the 3-gallon container into the 5-gallon container

Let $y \geq 0$ be the number of times we empty the 5-gallon container

Then, the amount of water in the 5-gallon container is $\ell = 3x - 5y$, a linear combination…



USING BEZOUT'S IDENTITY TO AVOID DISASTER

What does Bezout's Identity have to do with this?

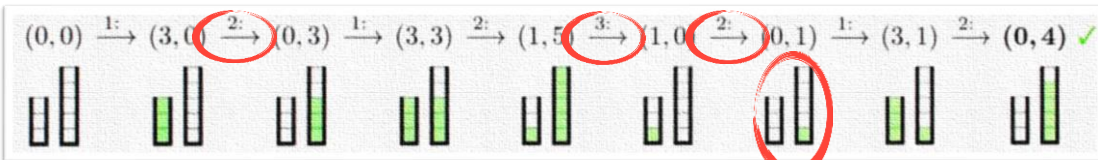
The amount of water in the 5-gallon container is $\ell = 3x - 5y$, a linear combination…

From Bezout's Identity, this linear combination implies $\ell = \gcd(3, 5)$, so we can get 1 gallon

$$\ell = 3x - 5y = 3 \times 2 - 5 \times 1 = 1$$

Here, we empty the 3-gallon container $x = 2$ times and the 5-gallon container $y = 1$ time

Repeat this four times and we have four gallons!



USING BEZOUT'S IDENTITY TO AVOID DISASTER

What does Bezout's Identity have to do with this?

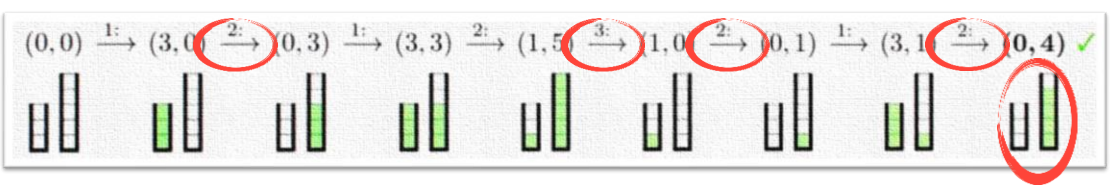
The amount of water in the 5-gallon container is $\ell = 3x - 5y$, a linear combination…

From Bezout's Identity, we want $\ell = 4$…

$$\ell = 4 = 3x - 5y = 3 \times 3 - 5 \times 1 = 4$$

Here, we empty the 3-gallon container $x = 3$ times and the 5-gallon container $y = 1$ time

We at least have shown we cannot do better than this solution…



WHAT NEXT…?

Problem Set 5 is due in your October 26 recitations

Watch for Homework 4 to be posted in Submitty later this week…

- …due by 11:59PM on Thursday, November 3

Practice! Tinker! Practice! Tinker! Practice! Tinker! Practice! Tinker! Practice!