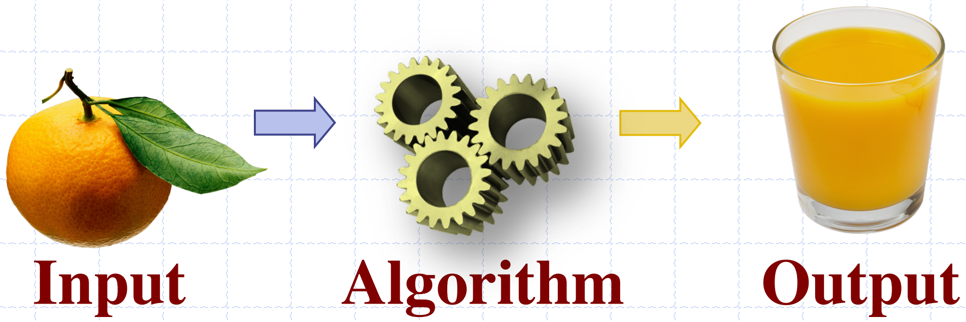
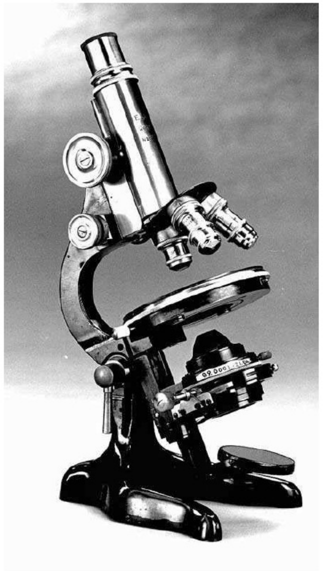
Presentation for use with the textbook, Algorithm Design and Applications, by M. T. Goodrich and R. Tamassia, Wiley, 2015

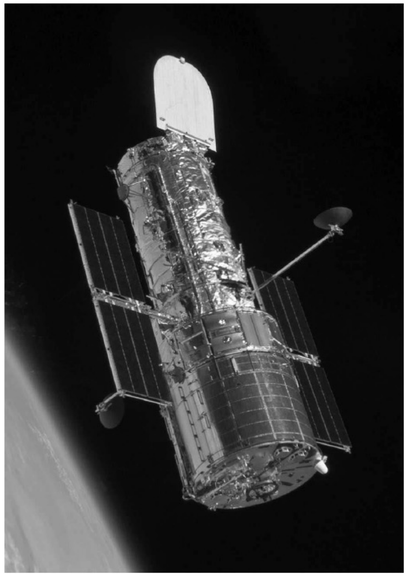
Analysis of Algorithms



Scalability

- Scientists often have to deal with differences in scale, from the microscopically small to the astronomically large.
- Computer scientists must also deal with scale, but they deal with it primarily in terms of data volume rather than physical object size.
- **Scalability** refers to the ability of a system to gracefully accommodate growing sizes of inputs or amounts of workload.

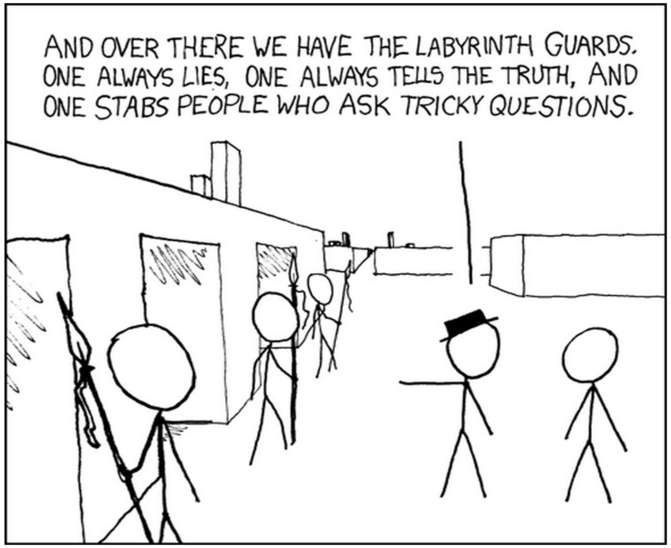




Microscope: U.S. government image, from the N.I.H. Medical Instrument Gallery, DeWitt Stetten, Jr., Museum of Medical Research. Hubble Space Telescope: U.S. government image, from NASA, STS-125 Crew, May 25, 2009.

Application: Job Interviews

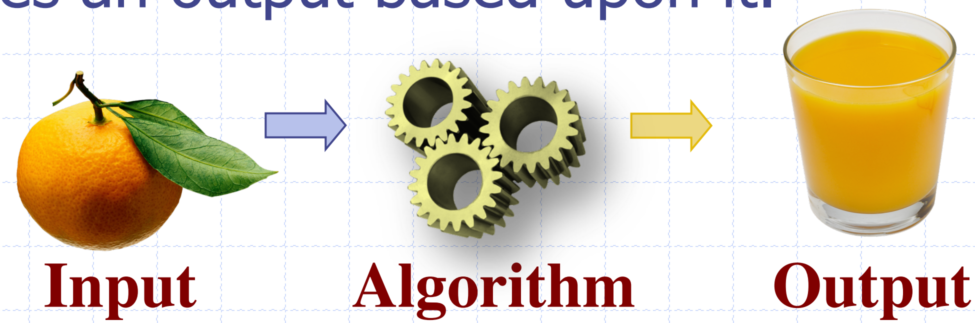
- High technology companies tend to ask questions about **algorithms and data structures** during job interviews.
- Algorithms questions can be short but often require critical thinking, creative insights, and subject knowledge.
  - All the "Applications" exercises in Chapter 1 of the Goodrich-Tamassia textbook are taken from reports of actual job interview questions.



xkcd "Labyrinth Puzzle." http://xkcd.com/246/

Algorithms and Data Structures

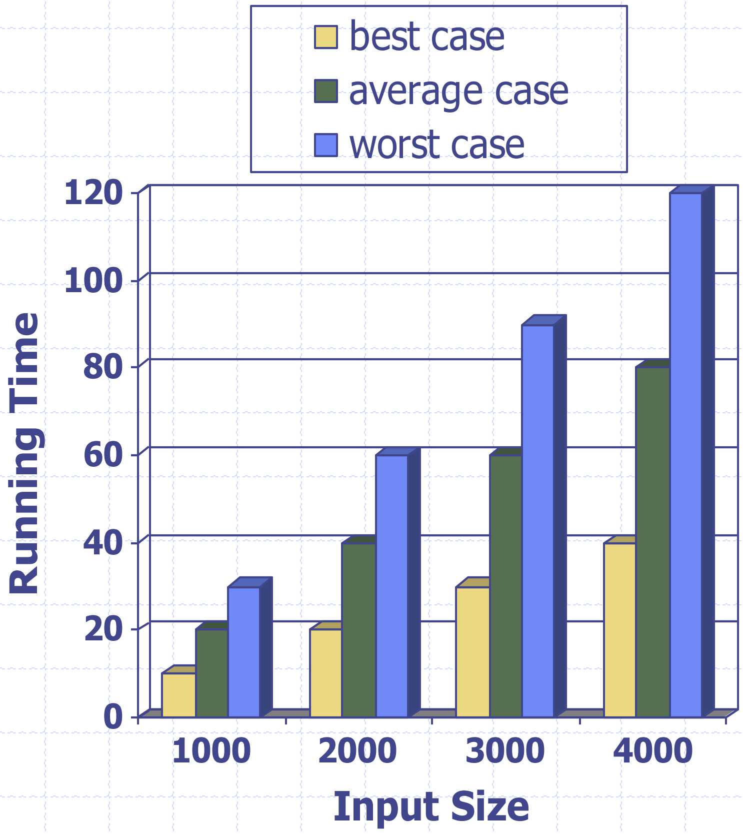
- An **algorithm** is a step-by-step procedure for performing some task in a finite amount of time.
  - Typically, an algorithm takes input data and produces an output based upon it.



- A **data structure** is a systematic way of organizing and accessing data.

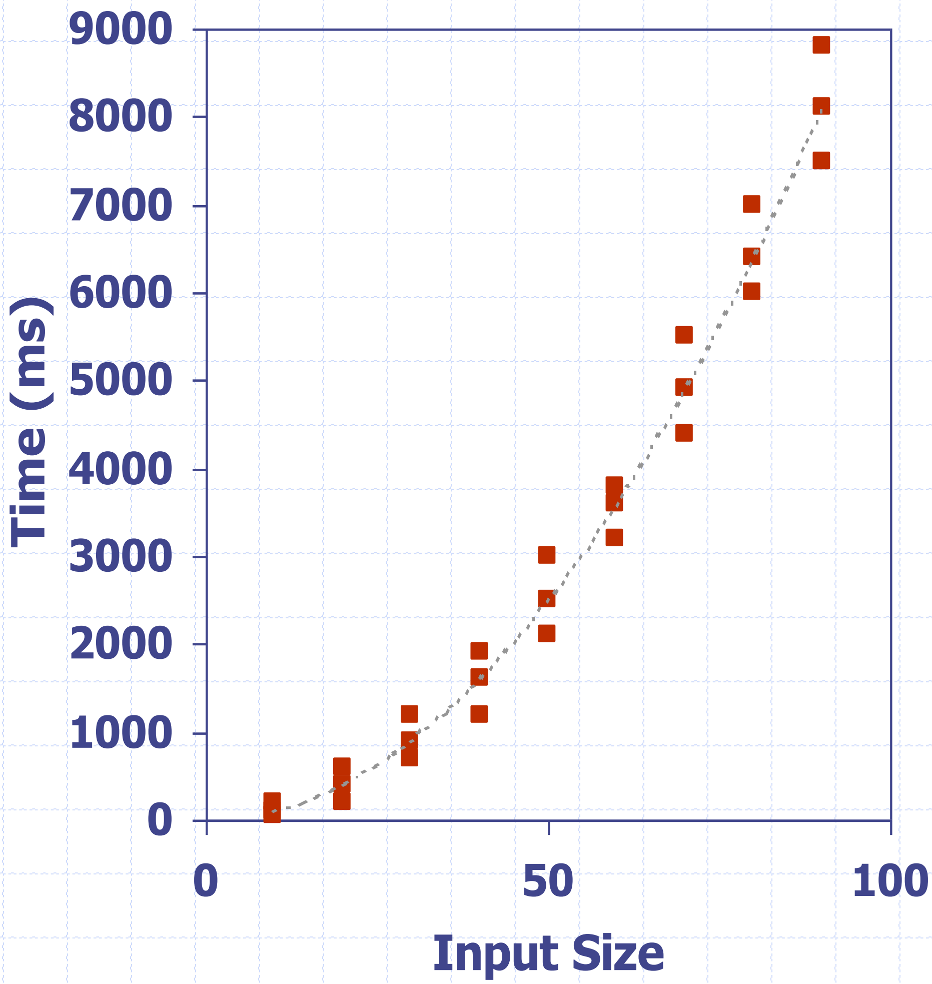
Running Time

- Most algorithms transform input objects into output objects.
- The running time of an algorithm typically grows with the input size.
- Average case time is often difficult to determine.
- We focus primarily on the **worst case running time**.
  - Easier to analyze
  - Crucial to applications such as games, finance and robotics



Experimental Studies

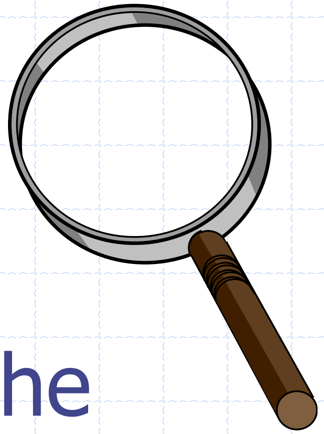
- Write a program implementing the algorithm
- Run the program with inputs of varying size and composition, noting the time needed:
- Plot the results



Limitations of Experiments

- It is necessary to implement the algorithm, which may be difficult
- Results may not be indicative of the running time on other inputs not included in the experiment.
- In order to compare two algorithms, the same hardware and software environments must be used

Theoretical Analysis



- Uses a high-level description of the algorithm instead of an implementation
- Characterizes running time as a function of the input size, n
- Takes into account all possible inputs
- Allows us to evaluate the speed of an algorithm independent of the hardware/software environment

Pseudocode

- High-level description of an algorithm
- More structured than English prose
- Less detailed than a program
- Preferred notation for describing algorithms
- Hides program design issues

Pseudocode Details

- Control flow
  - **if** … **then** … [**else** …]
  - **while** … **do** …
  - **repeat** … **until** …
  - **for** … **do** …
  - Indentation replaces braces
- Method declaration

  Algorithm *method* (*arg* [, *arg*…])

  Input …

  Output …
- Method call

  method (*arg* [, *arg*…])
- Return value

  return *expression*
- Expressions:

  $\leftarrow$ Assignment

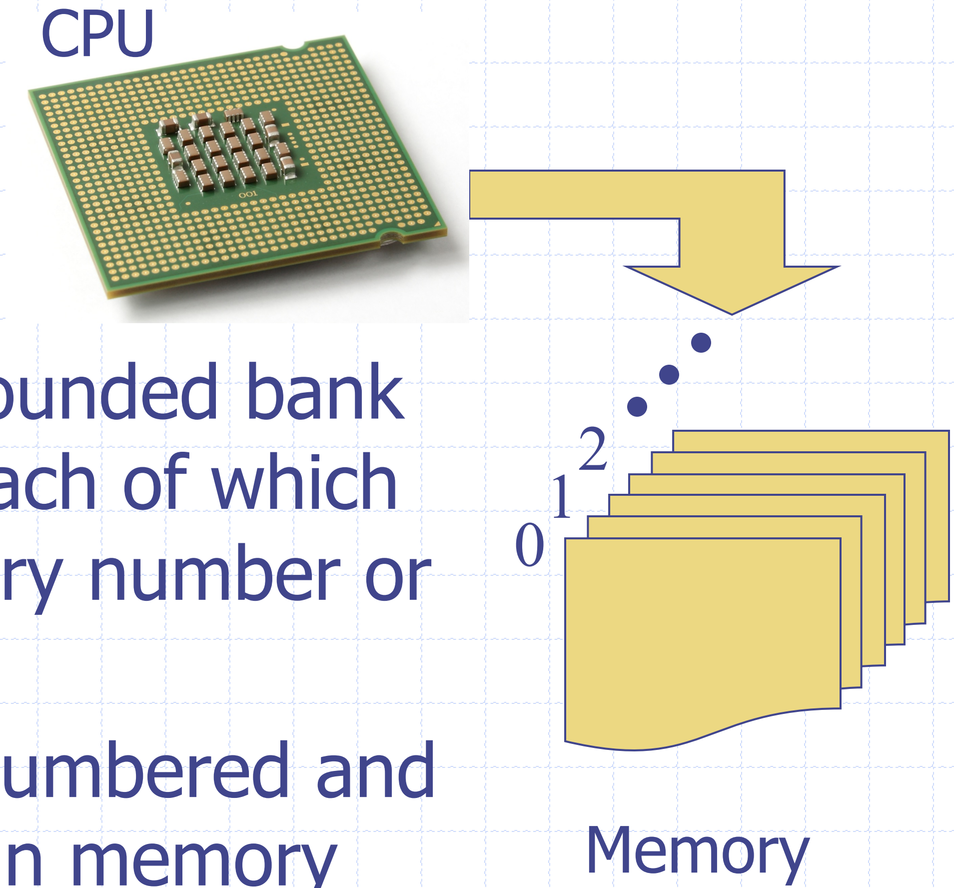
  $=$ Equality testing

  n^2 Superscripts and other mathematical formatting allowed

The Random Access Machine (RAM) Model

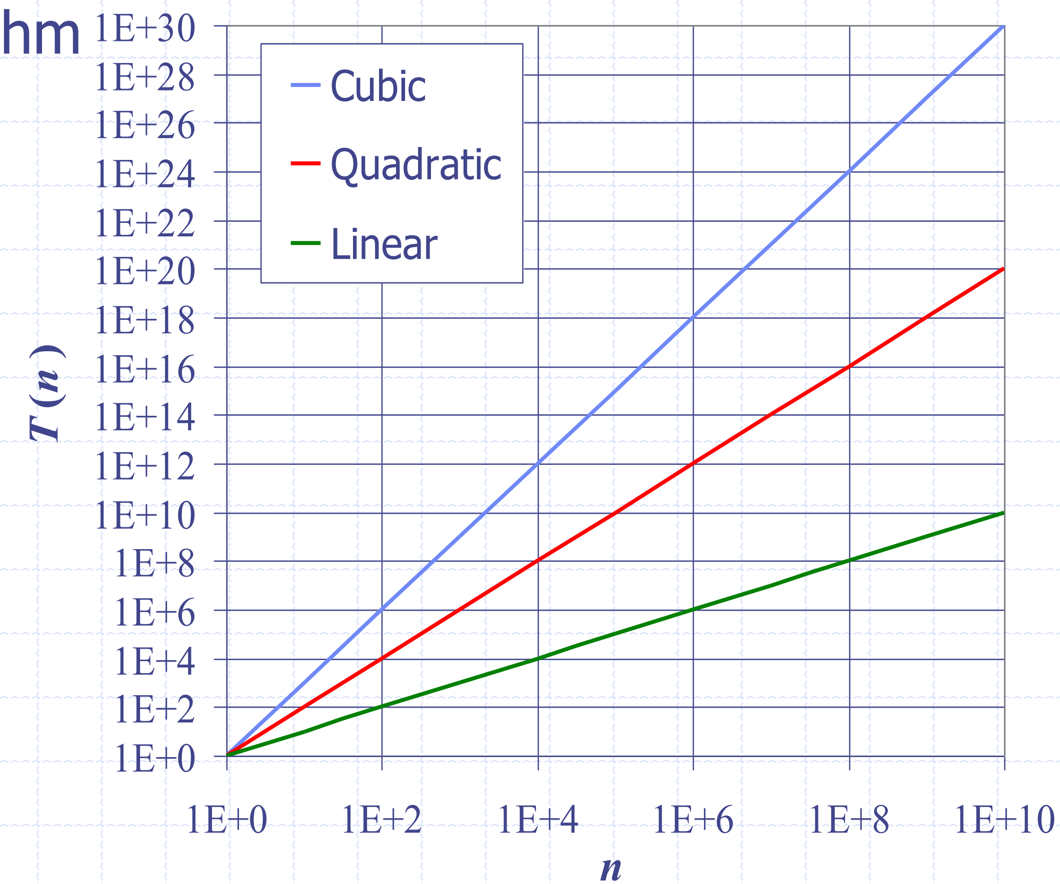
A **RAM** consists of

- A CPU
- An potentially unbounded bank of memory cells, each of which can hold an arbitrary number or character
- Memory cells are numbered and accessing any cell in memory takes unit time



Seven Important Functions

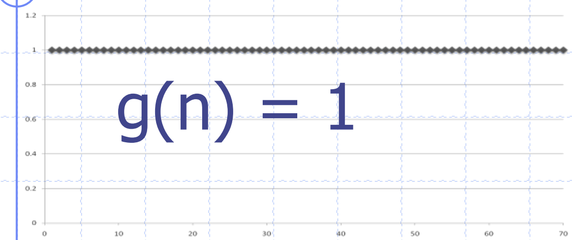
- Seven functions that often appear in algorithm analysis:
  - Constant ≈ 1
  - Logarithmic ≈ $\log n$
  - Linear ≈ n
  - N-Log-N ≈ $n \log n$
  - Quadratic ≈ n^2
  - Cubic ≈ n^3
  - Exponential ≈ 2^n

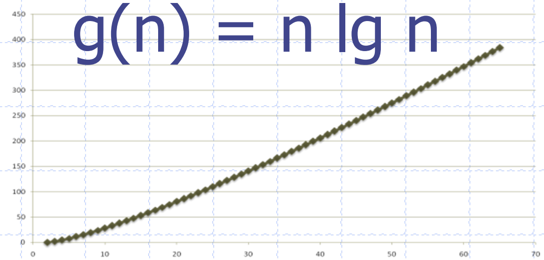


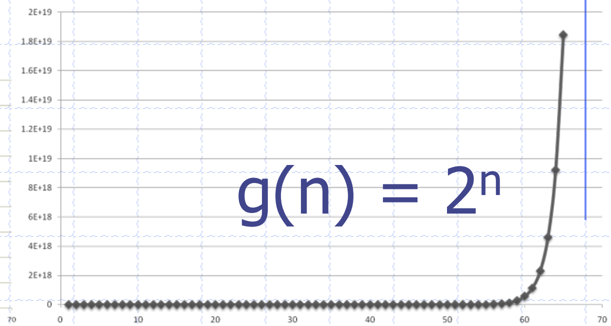
- In a log-log chart, the slope of the line corresponds to the growth rate

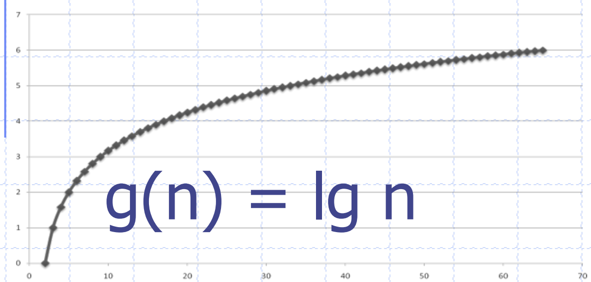
Functions Graphed Using “Normal” Scale

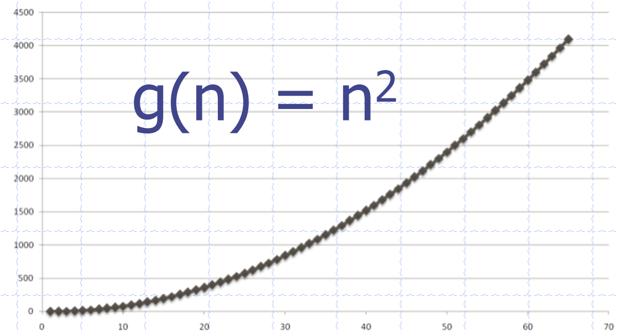
Slide by Matt Stallmann included with permission.

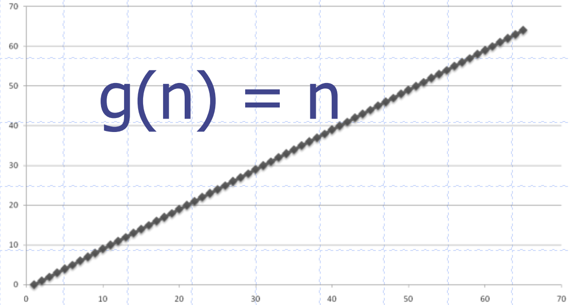


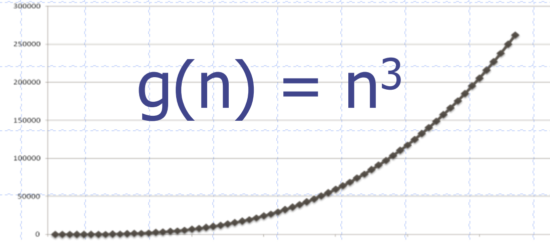












Primitive Operations

- Basic computations performed by an algorithm
- Identifiable in pseudocode
- Largely independent from the programming language
- Exact definition not important (we will see why later)
- Assumed to take a constant amount of time in the RAM model

- Examples:
  - Evaluating an expression
  - Assigning a value to a variable
  - Indexing into an array
  - Calling a method
  - Returning from a method

Counting Primitive Operations

- Example: By inspecting the pseudocode, we can determine the maximum number of primitive operations executed by an algorithm, as a function of the input size

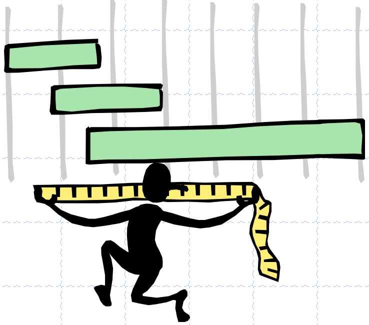
Algorithm arrayMax(A, n):

  Input: An array A storing $n \geq 1$ integers.

  Output: The maximum element in A.

  currentMax $\leftarrow A[0]$
  for $i \leftarrow 1$ **to** $n - 1$ **do**
    if *currentMax* $< A[i]$ **then**
      currentMax $\leftarrow A[i]$
  return *currentMax*

Estimating Running Time



- Algorithm arrayMax executes $7n - 2$ primitive operations in the worst case, $5n$ in the best case. Define:
  - a = Time taken by the fastest primitive operation
  - b = Time taken by the slowest primitive operation
- Let $T(n)$ be worst-case time of arrayMax. Then
$$a(5n) \le T(n) \le b(7n - 2)$$
- Hence, the running time $T(n)$ is bounded by two linear functions

Growth Rate of Running Time

- Changing the hardware/ software environment
  - Affects $T(n)$ by a constant factor, but
  - Does not alter the growth rate of $T(n)$
- The linear growth rate of the running time $T(n)$ is an intrinsic property of algorithm arrayMax

Slide by Matt Stallmann included with permission.

Why Growth Rate Matters

if runtime is...	time for n + 1	time for 2 n	time for 4 n
$c \lg n$	$c \lg (n + 1)$	$c (\lg n + 1)$	$c(\lg n + 2)$
$c\ n$	$c\ (n + 1)$	$2c\ n$	$4c\ n$
$c\ n \lg n$	$\sim c\ n \lg n + c\ n$	$2c\ n \lg n + 2cn$	$4c\ n \lg n + 4cn$
$c\ n^2$	$\sim c\ n^2 + 2c\ n$	**$4c\ n^2$**	$16c\ n^2$
$c\ n^3$	$\sim c\ n^3 + 3c\ n^2$	$8c\ n^3$	$64c\ n^3$
$c\ 2^n$	$c\ 2^{n+1}$	$c\ 2^{2n}$	$c\ 2^{4n}$

runtime quadruples when problem size doubles

Analyzing Recursive Algorithms

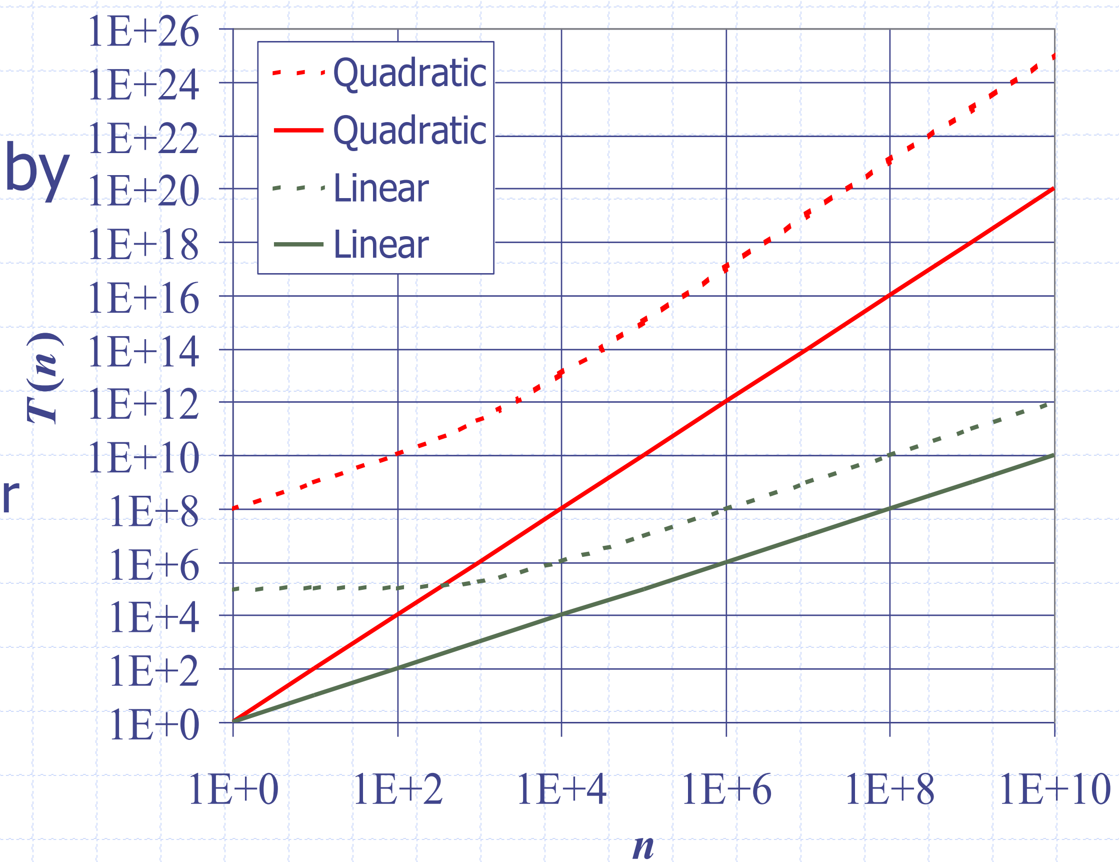
- Use a function, T(n), to derive a **recurrence relation** that characterizes the running time of the algorithm in terms of smaller values of n.

Algorithm recursiveMax(A, n):
    Input: An array A storing $n \geq 1$ integers.
    Output: The maximum element in A.
  if $n = 1$ **then**
      return $A[0]$
  return $\max\{\text{recursiveMax}(A, n-1),\ A[n-1]\}$

$$T(n) = \begin{cases} 3 & \text{if } n = 1 \\ T(n-1) + 7 & \text{otherwise,} \end{cases}$$

Constant Factors

- The growth rate is minimally affected by
  - constant factors or
  - lower-order terms
- Examples
  - $10^2 n + 10^5$ is a linear function
  - $10^5 n^2 + 10^8 n$ is a quadratic function

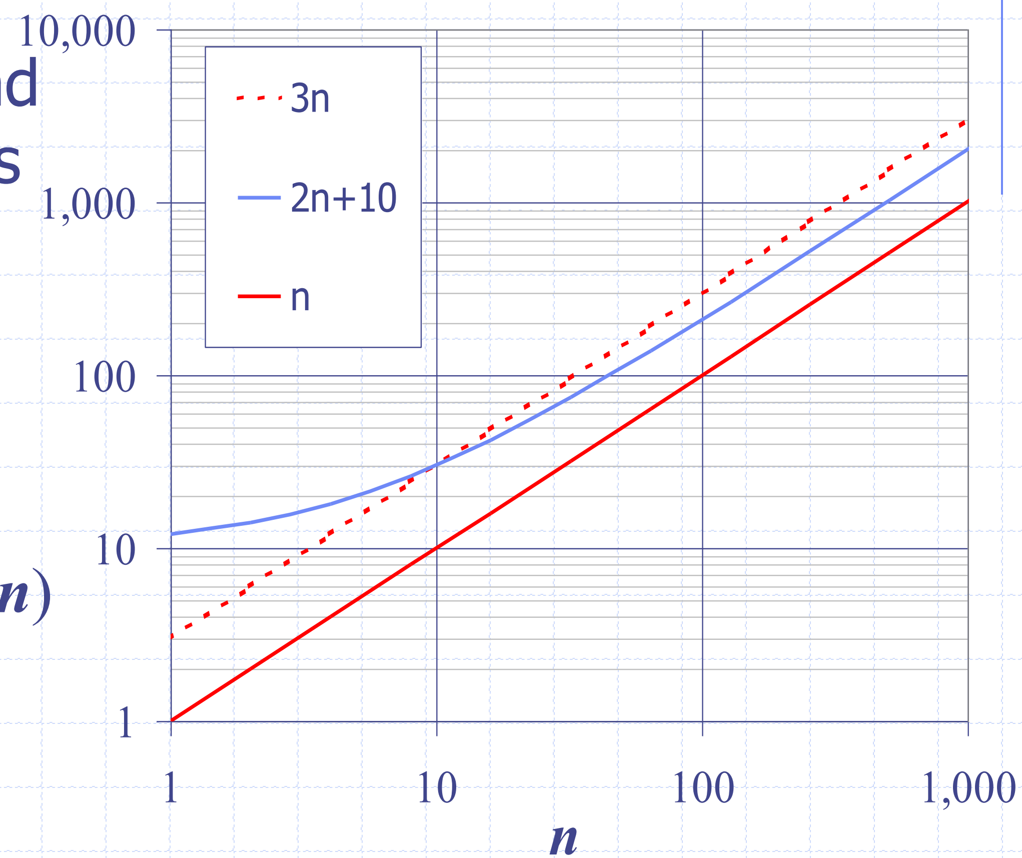


Big-Oh Notation

- Given functions $f(n)$ and $g(n)$, we say that $f(n)$ is $O(g(n))$ if there are positive constants c and n_0 such that

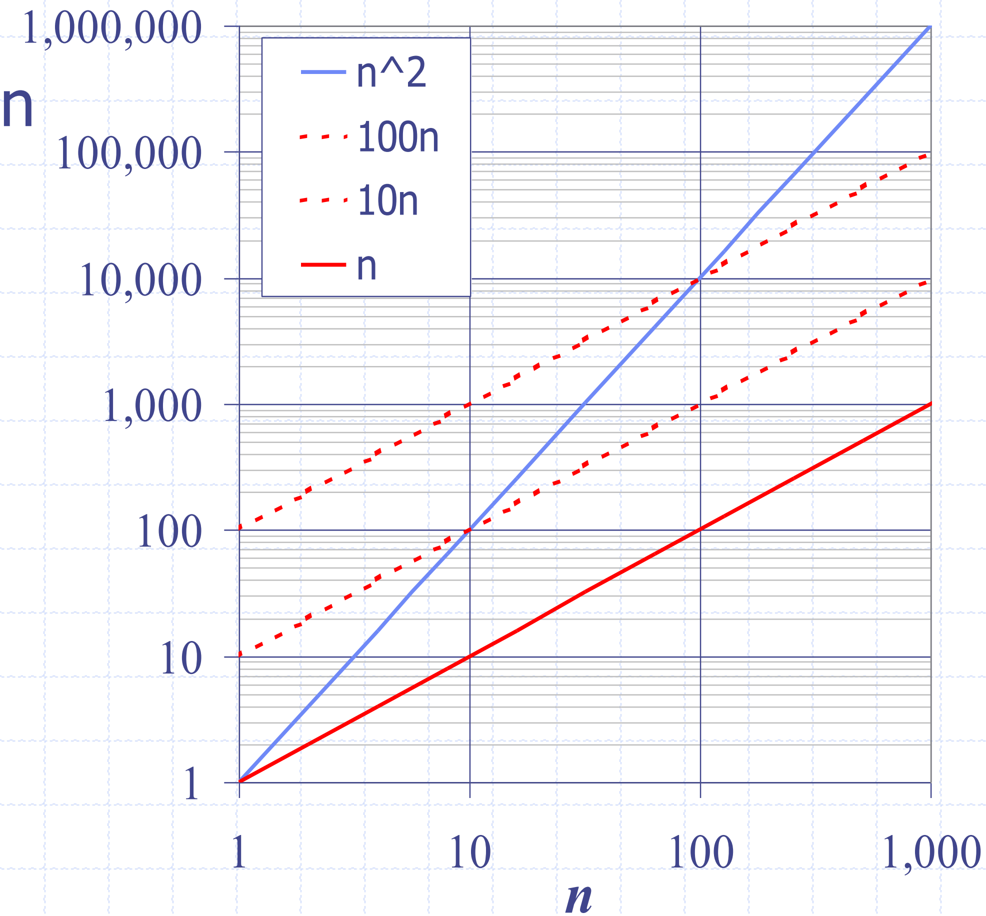
  $f(n) \leq cg(n)$ for $n \geq n_0$

- Example: $2n + 10$ is $O(n)$
  - $2n + 10 \leq cn$
  - $(c - 2)\, n \geq 10$
  - $n \geq 10/(c - 2)$
  - Pick $c = 3$ and $n_0 = 10$

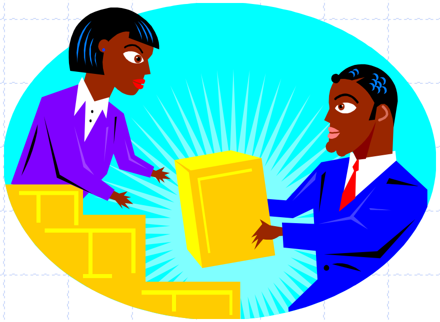


Big-Oh Example

- Example: the function n^2 is not $O(n)$
  - $n^2 \leq cn$
  - $n \leq c$
  - The above inequality cannot be satisfied since c must be a constant



More Big-Oh Examples



- $7n - 2$

  $7n-2$ is $O(n)$

  need $c > 0$ and $n_0 \geq 1$ such that $7n - 2 \leq cn$ for $n \geq n_0$

  this is true for $c = 7$ and $n_0 = 1$

- $3n^3 + 20n^2 + 5$

  $3n^3 + 20n^2 + 5$ is $O(n^3)$

  need $c > 0$ and $n_0 \geq 1$ such that $3n^3 + 20n^2 + 5 \leq cn^3$ for $n \geq n_0$

  this is true for $c = 4$ and $n_0 = 21$

- $3 \log n + 5$

  $3 \log n + 5$ is $O(\log n)$

  need $c > 0$ and $n_0 \geq 1$ such that $3 \log n + 5 \leq c \log n$ for $n \geq n_0$

  this is true for $c = 8$ and $n_0 = 2$

Big-Oh and Growth Rate

- The big-Oh notation gives an upper bound on the growth rate of a function
- The statement "$f(n)$ is $O(g(n))$" means that the growth rate of $f(n)$ is no more than the growth rate of $g(n)$
- We can use the big-Oh notation to rank functions according to their growth rate

	$f(n)$ is $O(g(n))$	$g(n)$ is $O(f(n))$
$g(n)$ grows more	Yes	No
$f(n)$ grows more	No	Yes
Same growth	Yes	Yes

Big-Oh Rules

- If is $f(n)$ a polynomial of degree d, then $f(n)$ is $O(n^d)$, i.e.,
  1. Drop lower-order terms
  2. Drop constant factors
- Use the smallest possible class of functions
  - Say "$2n$ is $O(n)$" instead of "$2n$ is $O(n^2)$"
- Use the simplest expression of the class
  - Say "$3n + 5$ is $O(n)$" instead of "$3n + 5$ is $O(3n)$"

Asymptotic Algorithm Analysis

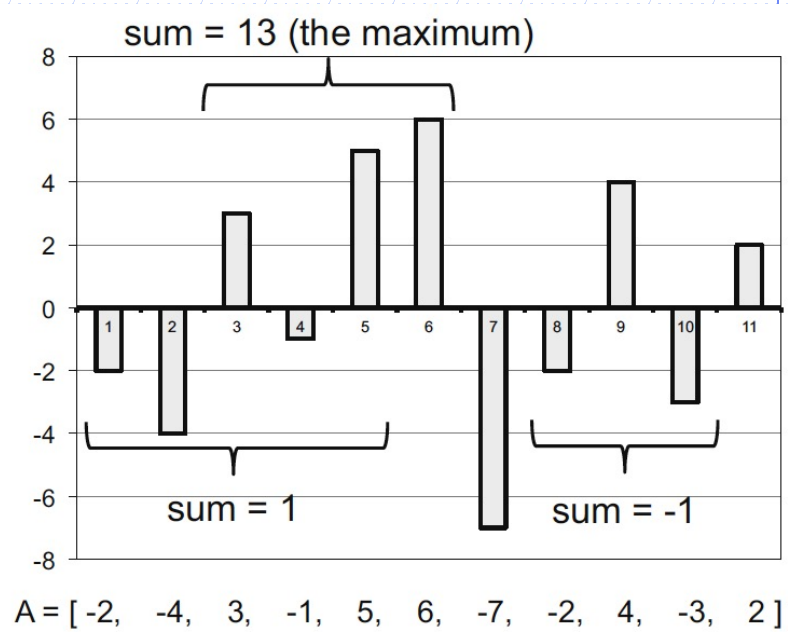
- The asymptotic analysis of an algorithm determines the running time in big-Oh notation
- To perform the asymptotic analysis
  - We find the worst-case number of primitive operations executed as a function of the input size
  - We express this function with big-Oh notation
- Example:
  - We say that algorithm arrayMax "runs in $O(n)$ time"
- Since constant factors and lower-order terms are eventually dropped anyhow, we can disregard them when counting primitive operations

A Case Study in Algorithm Analysis

- Given an array of *n* integers, find the subarray, A[j:k] that maximizes the sum

$$s_{j,k} = a_j + a_{j+1} + \cdots + a_k = \sum_{i=j}^{k} a_i.$$

- In addition to being an interview question for testing the thinking skills of job candidates, this **maximum subarray problem** also has applications in pattern analysis in digitized images.



A First (Slow) Solution

Compute the maximum of every possible subarray summation of the array *A* separately.

Algorithm MaxsubSlow(A):
   Input: An n-element array A of numbers, indexed from 1 to n.
   Output: The maximum subarray sum of array A.

```
m ← 0     // the maximum found so far
for j ← 1 to n do
    for k ← j to n do
        s ← 0     // the next partial sum we are computing
        for i ← j to k do
            s ← s + A[i]
        if s > m then
            m ← s
return m
```

- The outer loop, for index j, will iterate n times, its inner loop, for index k, will iterate at most n times, and the inner-most loop, for index i, will iterate at most n times.
- Thus, the running time of the MaxsubSlow algorithm is $O(n^3)$.

An Improved Algorithm

- A more efficient way to calculate these summations is to consider **prefix sums**

$$S_t = a_1 + a_2 + \cdots + a_t = \sum_{i=1}^{t} a_i$$

- If we are given all such prefix sums (and assuming $S_0=0$), we can compute any summation $s_{j,k}$ in constant time as

$$s_{j,k} = S_k - S_{j-1}$$

An Improved Algorithm, cont.

- Compute all the prefix sums
- Then compute all the subarray sums

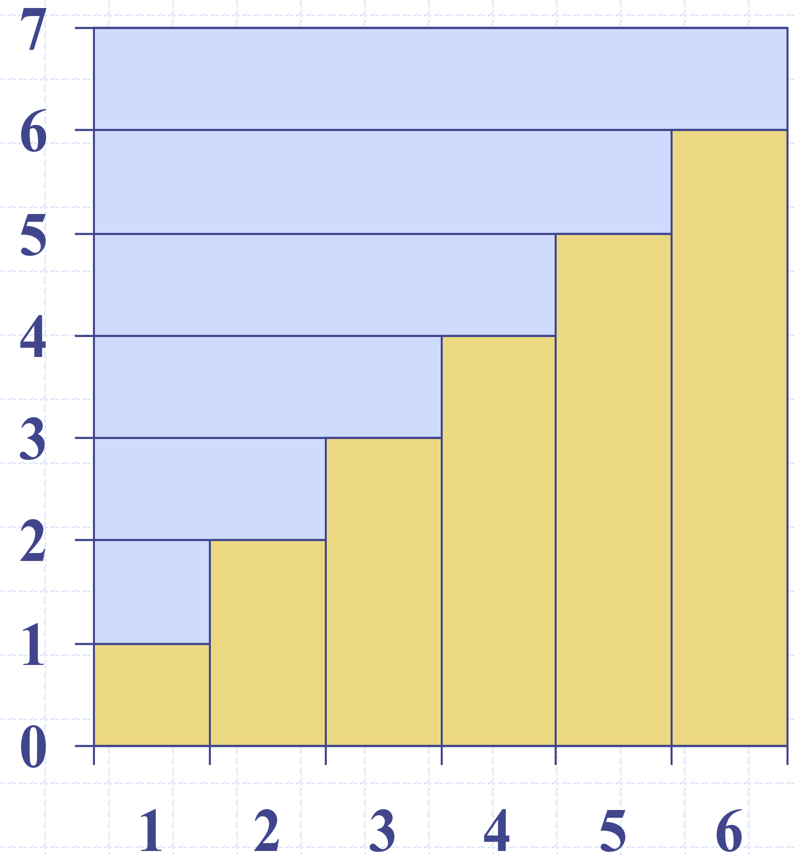
Algorithm MaxsubFaster(A):

  Input: An n-element array A of numbers, indexed from 1 to n.
  Output: The maximum subarray sum of array A.

```
S_0 ← 0     // the initial prefix sum
for i ← 1 to n do
    S_i ← S_{i−1} + A[i]
m ← 0     // the maximum found so far
for j ← 1 to n do
    for k ← j to n do
        s = S_k − S_{j−1}
        if s > m then
            m ← s
return m
```

Arithmetic Progression

- The running time of MaxsubFaster is $O(1 + 2 + \ldots + n)$
- The sum of the first n integers is $n(n + 1) / 2$
  - There is a simple visual proof of this fact
- Thus, algorithm MaxsubFaster runs in $O(n^2)$ time



A Linear-Time Algorithm

- Instead of computing prefix sum $S_t = s_{1,t}$, let us compute a maximum suffix sum, M_t, which is the maximum of 0 and the maximum $s_{j,t}$ for $j = 1,\ldots, t$.

$$M_t = \max\{0, \max_{j=1,\cdots,t}\{s_{j,t}\}\}$$

- if $M_t > 0$, then it is the summation value for a maximum subarray that ends at t, and if $M_t = 0$, then we can safely ignore any subarray that ends at t.
- if we know all the M_t values, for $t = 1, 2, \ldots, n$, then the solution to the maximum subarray problem would simply be the maximum of all these values.

A Linear-Time Algorithm, cont.

- for $t \geq 2$, if we have a maximum subarray that ends at t, and it has a positive sum, then it is either A[t : t] or it is made up of the maximum subarray that ends at $t - 1$ plus A[t]. So we can define $M_0 = 0$ and

$$M_t = \max\{0,\ M_{t-1} + A[t]\}$$

- If this were not the case, then we could make a subarray of even larger sum by swapping out the one we chose to end at $t - 1$ with the maximum one that ends at $t - 1$, which would contradict the fact that we have the maximum subarray that ends at t.
- Also, if taking the value of maximum subarray that ends at $t - 1$ and adding A[t] makes this sum no longer be positive, then $M_t = 0$, for there is no subarray that ends at t with a positive summation.

A Linear-Time Algorithm, cont.

Algorithm MaxsubFastest(A):

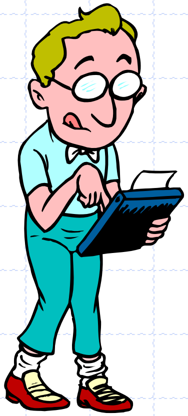
Input: An n-element array A of numbers, indexed from 1 to n.

Output: The maximum subarray sum of array A.

$M_0 \leftarrow 0$ // the initial prefix maximum
for $t \leftarrow 1$ **to** n **do**
  $M_t \leftarrow \max\{0, M_{t-1} + A[t]\}$
$m \leftarrow 0$ // the maximum found so far
for $t \leftarrow 1$ **to** n **do**
  $m \leftarrow \max\{m, M_t\}$
return m

- The MaxsubFastest algorithm consists of two loops, which each iterate exactly n times and take O(1) time in each iteration. Thus, the total running time of the MaxsubFastest algorithm is O(n).

Math you need to Review



- Summations
- Powers
- Logarithms
- Proof techniques
- Basic probability

- Properties of powers:
  $a^{(b+c)} = a^b a^c$
  $a^{bc} = (a^b)^c$
  $a^b / a^c = a^{(b-c)}$
  $b = a^{\log_a b}$
  $b^c = a^{c*\log_a b}$
- Properties of logarithms:
  $\log_b(xy) = \log_b x + \log_b y$
  $\log_b (x/y) = \log_b x - \log_b y$
  $\log_b xa = a\log_b x$
  $\log_b a = \log_x a / \log_x b$

Relatives of Big-Oh

big-Omega

- f(n) is Ω(g(n)) if there is a constant c > 0 and an integer constant $n_0 \geq 1$ such that

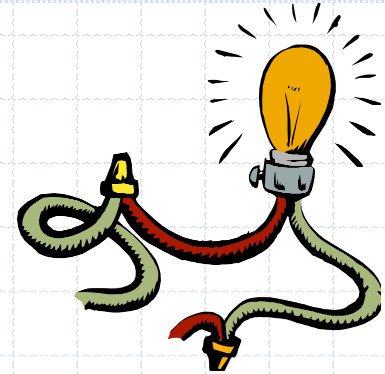
$$f(n) \geq c\, g(n) \text{ for } n \geq n_0$$

big-Theta

- f(n) is Θ(g(n)) if there are constants c′ > 0 and c″ > 0 and an integer constant $n_0 \geq 1$ such that

$$c'g(n) \leq f(n) \leq c''g(n) \text{ for } n \geq n_0$$

Intuition for Asymptotic Notation



big-Oh

- f(n) is O(g(n)) if f(n) is asymptotically less than or equal to g(n)

big-Omega

- f(n) is Ω(g(n)) if f(n) is asymptotically greater than or equal to g(n)

big-Theta

- f(n) is Θ(g(n)) if f(n) is asymptotically equal to g(n)

Example Uses of the Relatives of Big-Oh

- **$5n^2$ is $\Omega(n^2)$**

  $f(n)$ is $\Omega(g(n))$ if there is a constant $c > 0$ and an integer constant $n_0 \geq 1$ such that $f(n) \geq c\, g(n)$ for $n \geq n_0$

  let $c = 5$ and $n_0 = 1$

- **$5n^2$ is $\Omega(n)$**

  $f(n)$ is $\Omega(g(n))$ if there is a constant $c > 0$ and an integer constant $n_0 \geq 1$ such that $\mathrm{f}(n) \geq c\, g(n)$ for $n \geq n_0$

  let $c = 1$ and $n_0 = 1$

- **$5n^2$ is $\Theta(n^2)$**

  $f(n)$ is $\Theta(g(n))$ if it is $\Omega(n^2)$ and $O(n^2)$. We have already seen the former, for the latter recall that $f(n)$ is $O(g(n))$ if there is a constant $c > 0$ and an integer constant $n_0 \geq 1$ such that $\mathrm{f}(n) \leq c\, g(n)$ for $n \geq n_0$

  Let $c = 5$ and $n_0 = 1$

O ("big oh")

O ("big oh")

Informally:

O ("big oh")

Informally:

- $g \in O(f)$ if g is bounded above by a constant multiple of f (for sufficiently large values of n).

O ("big oh")

Informally:

- $g \in O(f)$ if g is bounded above by a constant multiple of f (for sufficiently large values of n).
- $g \in O(f)$ if "g grows no faster than (a constant multiple of) f."

O ("big oh")

Informally:

- $g \in O(f)$ if g is bounded above by a constant multiple of f (for sufficiently large values of n).
- $g \in O(f)$ if "g grows no faster than (a constant multiple of) f."
- $g \in O(f)$ if the ratio g/f is bounded above by a constant (for sufficiently values of n).

O ("big oh")

O ("big oh")

Formally:

O ("big oh")

Formally:

- $g \in O(f)$ if and only if:

$$\exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, g(n) \leq C \cdot f(n).$$

O ("big oh")

Formally:

- $g \in O(f)$ if and only if:

$$\exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, g(n) \leq C \cdot f(n).$$

- Equivalently: $g \in O(f)$ if and only if:

$$\exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, \frac{g(n)}{f(n)} \leq C.$$

O ("big oh")

Formally:

- $g \in O(f)$ if and only if:

$$\exists_{C>0} \, \exists_{n_0>0} \, \forall_{n>n_0} \, g(n) \leq C \cdot f(n).$$

- Equivalently: $g \in O(f)$ if and only if:

$$\exists_{C>0} \, \exists_{n_0>0} \, \forall_{n>n_0} \, \frac{g(n)}{f(n)} \leq C.$$

- Sometimes we write: $g = O(f)$ rather than $g \in O(f)$

Examples of O-notation:

Example 1: $f(n) = n$, $g(n) = 1000n$: $g \in O(f)$.

Examples of O-notation:

Example 1: $f(n) = n$, $g(n) = 1000n$: $g \in O(f)$.

Proof:

Examples of O-notation:

Example 1: $f(n) = n$, $g(n) = 1000n$: $g \in O(f)$.

Proof: Let $C = 1000$. Then $g(n) \leq C \cdot f(n)$ for all n.

Examples of O-notation:

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof:

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.
Hence for any $C > 0$ the ratio is less than C as long as n is sufficiently large.

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.
Hence for any $C > 0$ the ratio is less than C as long as n is sufficiently large.(Of course, how large n must be to be"sufficiently large" depends on C).

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.
Hence for any $C > 0$ the ratio is less than C as long as n is sufficiently large.(Of course, how large n must be to be"sufficiently large" depends on C).

Alternate Proof:

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.
Hence for any $C > 0$ the ratio is less than C as long as n is sufficiently large.(Of course, how large n must be to be"sufficiently large" depends on C).

Alternate Proof: If $n \geq 1$, $n^{1/2} \geq 1$, so $n^{3/2} \leq n^2$.

Examples of O-notation:

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$: $g \in O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.
Hence for any $C > 0$ the ratio is less than C as long as n is sufficiently large.(Of course, how large n must be to be"sufficiently large" depends on C).

Alternate Proof: If $n \geq 1$, $n^{1/2} \geq 1$, so $n^{3/2} \leq n^2$.
Hence we can choose $C = 1$ and $n_0 = 1$.

Examples of O-notation:

Examples of O-notation:

Example 3: $f(n) = n^3$, $g(n) = n^4$: $g \notin O(f)$.

Examples of O-notation:

Example 3: $f(n) = n^3$, $g(n) = n^4$: $g \notin O(f)$.

Proof:

Examples of O-notation:

Example 3: $f(n) = n^3$, $g(n) = n^4$: $g \notin O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = \infty$.

Examples of O-notation:

Example 3: $f(n) = n^3$, $g(n) = n^4$: $g \notin O(f)$.

Proof: $\lim_{n\to\infty} \frac{g(n)}{f(n)} = \infty$.
Hence there is no $C > 0$ such that $g(n) \leq C \cdot f(n)$ for sufficiently large n.

Examples of O-notation:

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof:

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$.

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$. Hence:

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$. Hence:

$$g(n) \;=\; 5n^2 + 23n + 2$$

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$. Hence:

$$\begin{aligned} g(n) &= 5n^2 + 23n + 2 \\ &\leq 5n^2 + 23n^2 + 2n^2 \end{aligned}$$

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$. Hence:

$$\begin{aligned} g(n) &= 5n^2 + 23n + 2 \\ &\leq 5n^2 + 23n^2 + 2n^2 \\ &\leq 30n^2 \end{aligned}$$

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$. Hence:

$$\begin{aligned} g(n) &= 5n^2 + 23n + 2 \\ &\leq 5n^2 + 23n^2 + 2n^2 \\ &\leq 30n^2 \\ &= 30f(n) \end{aligned}$$

Examples of O-notation:

Example 4: $f(n) = n^2$, $g(n) = 5n^2 + 23n + 2$: $g \in O(f)$.

Proof: If $n \geq 1$, then $n \leq n^2$ and $1 \leq n^2$. Hence:

$$\begin{aligned} g(n) &= 5n^2 + 23n + 2 \\ &\leq 5n^2 + 23n^2 + 2n^2 \\ &\leq 30n^2 \\ &= 30f(n) \end{aligned}$$

So we can take $C = 30$, $n_0 = 1$.

More asymptotic notation: o ("little oh"), Ω ("big Omega")

More asymptotic notation: o ("little oh"), Ω ("big Omega")

- o ('little oh"):

More asymptotic notation: o (“little oh”), Ω (“big Omega”)

- o (‘little oh”):

$$g \in o(f) \quad \text{if and only if} \lim_{n\to\infty} \frac{g(n)}{f(n)} = 0.$$

More asymptotic notation: o ("little oh"), Ω ("big Omega")

- o ('little oh"):

$$g \in o(f) \quad \text{if and only if } \lim_{n \to \infty} \frac{g(n)}{f(n)} = 0.$$

- Ω ("big Omega") (or just "Omega")

More asymptotic notation: o ("little oh"), Ω ("big Omega")

- o ('little oh"):

$$g \in o(f) \quad \text{if and only if } \lim_{n\to\infty} \frac{g(n)}{f(n)} = 0.$$

- Ω ("big Omega") (or just "Omega")

$$g \in \Omega(f) \quad \text{if and only if} \quad \exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, g(n) \geq C \cdot f(n).$$

More asymptotic notation: o ("little oh"), Ω ("big Omega")

- o ('little oh"):

$$g \in o(f) \quad \text{if and only if} \lim_{n\to\infty} \frac{g(n)}{f(n)} = 0.$$

- Ω ("big Omega") (or just "Omega")

$$g \in \Omega(f) \quad \text{if and only if} \quad \exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, g(n) \geq C \cdot f(n).$$

Equivalently:

More asymptotic notation: o ("little oh"), Ω ("big Omega")

- o ('little oh"):

$$g \in o(f) \quad \text{if and only if} \lim_{n\to\infty} \frac{g(n)}{f(n)} = 0.$$

- Ω ("big Omega") (or just "Omega")

$$g \in \Omega(f) \quad \text{if and only if} \quad \exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, g(n) \geq C \cdot f(n).$$

Equivalently:

$$g \in \Omega(f) \quad \text{if and only if} \quad \exists_{C>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\, \frac{g(n)}{f(n)} \geq C.$$

One more definition: Θ ("Theta")

One more definition: Θ ("Theta")

- $g \in \Theta(f)$ if and only if:

$$g \in O(f) \text{ and } g \in \Omega(f).$$

One more definition: Θ ("Theta")

- $g \in \Theta(f)$ if and only if:

$$g \in O(f) \text{ and } g \in \Omega(f).$$

- Equivalently, $g \in \Theta(f)$ if and only if:

$$\exists_{C_1>0}\, \exists_{C_2>0}\, \exists_{n_0>0}\, \forall_{n>n_0}\; C_1 \leq \frac{g(n)}{f(n)} \leq C_2.$$

Examples of Asymptotic notation

Examples of Asymptotic notation

Example 1: $f(n) = n$, $g(n) = 1000n$.

Examples of Asymptotic notation

Example 1: $f(n) = n$, $g(n) = 1000n$.

$g \in \Omega(f)$, $g \in \Theta(f)$

Examples of Asymptotic notation

Example 1: $f(n) = n$, $g(n) = 1000n$.

$g \in \Omega(f)$, $g \in \Theta(f)$

To see that $g \in \Omega(f)$, we can take $C = 1$.

Examples of Asymptotic notation

Example 1: $f(n) = n$, $g(n) = 1000n$.

$g \in \Omega(f)$, $g \in \Theta(f)$

To see that $g \in \Omega(f)$, we can take $C = 1$.

Then $g(n) = 1000 \cdot n > 1 \cdot n = C \cdot f(n)$.

Examples of Asymptotic notation

Example 1: $f(n) = n$, $g(n) = 1000n$.

$g \in \Omega(f)$, $g \in \Theta(f)$

To see that $g \in \Omega(f)$, we can take $C = 1$.

Then $g(n) = 1000 \cdot n > 1 \cdot n = C \cdot f(n)$.

To see that $g \in \Theta(f)$, we could argue that $g \in O(f)$ (shown earlier) and $g \in \Omega(f)$ (shown above).

Examples of Asymptotic notation

Example 1: $f(n) = n$, $g(n) = 1000n$.

$g \in \Omega(f)$, $g \in \Theta(f)$

To see that $g \in \Omega(f)$, we can take $C = 1$.

Then $g(n) = 1000 \cdot n > 1 \cdot n = C \cdot f(n)$.

To see that $g \in \Theta(f)$, we could argue that $g \in O(f)$ (shown earlier) and $g \in \Omega(f)$ (shown above).

Or we can take $C_1 = 1$, $C_2 = 1000$. Then

$$C_1 \leq \frac{g(n)}{f(n} \leq C_2.$$

Examples of Asymptotic notation

Examples of Asymptotic notation

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$:

Examples of Asymptotic notation

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$:

$g \in o(f)$

Examples of Asymptotic notation

Example 2: $f(n) = n^2$, $g(n) = n^{3/2}$:

$g \in o(f)$

Because $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.

Examples of Asymptotic notation

Examples of Asymptotic notation

Example 3: $f(n) = n^3$, $g(n) = n^4$:

Examples of Asymptotic notation

Example 3: $f(n) = n^3$, $g(n) = n^4$:

$g \in \Omega(f)$

Examples of Asymptotic notation

Example 3: $f(n) = n^3$, $g(n) = n^4$:

$g \in \Omega(f)$

Because $\lim_{n\to\infty} \frac{g(n)}{f(n)} = \infty$, so we can choose any C we want.

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof:

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$.

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$. Hence if $n \geq 23$:

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$. Hence if $n \geq 23$:

$$g(n) = 5n^2 - 23n + 2$$

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$. Hence if $n \geq 23$:

$$\begin{aligned} g(n) &= 5n^2 - 23n + 2 \\ &\geq 5n^2 - n^2 \end{aligned}$$

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$. Hence if $n \geq 23$:

$$\begin{aligned} g(n) &= 5n^2 - 23n + 2 \\ &\geq 5n^2 - n^2 \\ &\geq 4n^2 \end{aligned}$$

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$. Hence if $n \geq 23$:

$$\begin{aligned} g(n) &= 5n^2 - 23n + 2 \\ &\geq 5n^2 - n^2 \\ &\geq 4n^2 \\ &= 4f(n) \end{aligned}$$

Examples of Asymptotic notation

Example 4: $f(n) = n^2$, $g(n) = 5n^2 - 23n + 2$:

$g \in \Omega(f)$.

Proof: If $n \geq 23$, then $23n \leq n^2$. Hence if $n \geq 23$:

$$\begin{aligned} g(n) &= 5n^2 - 23n + 2 \\ &\geq 5n^2 - n^2 \\ &\geq 4n^2 \\ &= 4f(n) \end{aligned}$$

So we can take $C = 4$, $n_0 = 23$.

Another Example

Another Example

Example 5: $\ln n = o(n)$

Another Example

Example 5: $\ln n = o(n)$

Proof:

Another Example

Example 5: $\ln n = o(n)$

Proof:

Examine the ratio $\frac{\ln n}{n}$ as $n \to \infty$.

Another Example

Example 5: $\ln n = o(n)$

Proof:

Examine the ratio $\frac{\ln n}{n}$ as $n \to \infty$.

If we try to evaluate the limit directly, we obtain the "indeterminate form" $\frac{\infty}{\infty}$.

Another Example

Example 5: $\ln n = o(n)$

Proof:

Examine the ratio $\frac{\ln n}{n}$ as $n \to \infty$.

If we try to evaluate the limit directly, we obtain the "indeterminate form" $\frac{\infty}{\infty}$.

We need to apply L'Hôpital's rule (from calculus).

Another Example

Example 5: $\ln n = o(n)$

Proof:

Examine the ratio $\frac{\ln n}{n}$ as $n \to \infty$.

If we try to evaluate the limit directly, we obtain the "indeterminate form" $\frac{\infty}{\infty}$.

We need to apply L'Hôpital's rule (from calculus).

(Continued on next slide)

Example 5, continued: $\ln n = o(n)$

L'Hôpital's rule: If the ratio of limits

$$\frac{\lim_{n\to\infty} g(n)}{\lim_{n\to\infty} f(n)}$$

is an indeterminate form (i.e., ∞/∞ or $0/0$), then

$$\lim_{n\to\infty} \frac{g(n)}{f(n)} = \lim_{n\to\infty} \frac{g'(n)}{f'(n)}$$

where f' and g' are, respectively, the derivatives of f and g.

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Example 5, continued:
$\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

By L'Hôpital's rule:

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

By L'Hôpital's rule:

$$\lim_{n\to\infty} \frac{g(n)}{f(n)} \quad = \quad \lim_{n\to\infty} \frac{g'(n)}{f'(n)}$$

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

By L'Hôpital's rule:

$$\begin{aligned} \lim_{n\to\infty} \frac{g(n)}{f(n)} &= \lim_{n\to\infty} \frac{g'(n)}{f'(n)} \\ &= \lim_{n\to\infty} \frac{1/n}{1} \end{aligned}$$

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

By L'Hôpital's rule:

$$\begin{aligned}
\lim_{n\to\infty} \frac{g(n)}{f(n)} &= \lim_{n\to\infty} \frac{g'(n)}{f'(n)} \\
&= \lim_{n\to\infty} \frac{1/n}{1} \\
&= \lim_{n\to\infty} \frac{1}{n}
\end{aligned}$$

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

By L'Hôpital's rule:

$$\begin{aligned}
\lim_{n\to\infty} \frac{g(n)}{f(n)} &= \lim_{n\to\infty} \frac{g'(n)}{f'(n)} \\
&= \lim_{n\to\infty} \frac{1/n}{1} \\
&= \lim_{n\to\infty} \frac{1}{n} \\
&= 0.
\end{aligned}$$

Example 5, continued: $\ln n = o(n)$

Let $f(n) = n$, $g(n) = \ln n$.

Then $f'(n) = 1$, $g'(n) = 1/n$.

By L'Hôpital's rule:

$$\begin{aligned}
\lim_{n\to\infty} \frac{g(n)}{f(n)} &= \lim_{n\to\infty} \frac{g'(n)}{f'(n)} \\
&= \lim_{n\to\infty} \frac{1/n}{1} \\
&= \lim_{n\to\infty} \frac{1}{n} \\
&= 0.
\end{aligned}$$

Hence $g(n) = o(f(n))$.

Math background

Math background

- Sums, Summations

Math background

- Sums, Summations
- Logarithms, Exponents Floors, Ceilings, Harmonic Numbers

Math background

- Sums, Summations
- Logarithms, Exponents Floors, Ceilings, Harmonic Numbers
- Proof Techniques

Math background

- Sums, Summations
- Logarithms, Exponents Floors, Ceilings, Harmonic Numbers
- Proof Techniques
- Basic Probability

Sums, Summations

Sums, Summations

- Summation notation:

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:
  - What if $a = b$?

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:
  - What if $a = b$? $f(a)$

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:
  - What if $a = b$? $f(a)$
  - What if $a > b$?

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:
  - What if $a = b$? $f(a)$
  - What if $a > b$? 0

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:
  - What if $a = b$? $f(a)$
  - What if $a > b$? 0
- If $S = \{s_1, \ldots, s_n\}$ is a finite set:

Sums, Summations

- Summation notation:

$$\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b).$$

- Special cases:
  - What if $a = b$? $f(a)$
  - What if $a > b$? 0
- If $S = \{s_1, \ldots, s_n\}$ is a finite set:

$$\sum_{x \in S} f(x) = f(s_1) + f(s_2) + \cdots + f(s_n).$$

Geometric sum

Geometric sum

- Geometric sum:

Geometric sum

- Geometric sum:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

provided that $a \neq 1$.

Geometric sum

- Geometric sum:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

  provided that $a \neq 1$.
- Previous formula holds for $a = 0$ because $a^0 = 1$ even when $a = 0$.

Geometric sum

- Geometric sum:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

  provided that $a \neq 1$.
- Previous formula holds for $a = 0$ because $a^0 = 1$ even when $a = 0$.
- Special case of geometric sum:

Geometric sum

- Geometric sum:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

  provided that $a \neq 1$.
- Previous formula holds for $a = 0$ because $a^0 = 1$ even when $a = 0$.
- Special case of geometric sum:

$$\sum_{i=0}^{n} 2^i = 1 + 2 + 4 + 8 + \cdots + 2^n = 2^{n+1} - 1.$$

Infinite Geometric sum

Infinite Geometric sum

- From the previous slide:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

provided that $a \neq 1$.

Infinite Geometric sum

- From the previous slide:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

provided that $a \neq 1$.

- If $|a| < 1$, we can take the limit as $n \to \infty$:

Infinite Geometric sum

- From the previous slide:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

  provided that $a \neq 1$.

- If $|a| < 1$, we can take the limit as $n \to \infty$:

$$\sum_{i=0}^{\infty} a^i = 1 + a^1 + a^2 + \cdots = \frac{1}{1 - a},$$

Infinite Geometric sum

- From the previous slide:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

provided that $a \neq 1$.

- If $|a| < 1$, we can take the limit as $n \to \infty$:

$$\sum_{i=0}^{\infty} a^i = 1 + a^1 + a^2 + \cdots = \frac{1}{1 - a},$$

- Special case of infinite geometric sum:

Infinite Geometric sum

- From the previous slide:

$$\sum_{i=0}^{n} a^i = 1 + a^1 + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a},$$

provided that $a \neq 1$.

- If $|a| < 1$, we can take the limit as $n \to \infty$:

$$\sum_{i=0}^{\infty} a^i = 1 + a^1 + a^2 + \cdots = \frac{1}{1 - a},$$

- Special case of infinite geometric sum:

$$\sum_{i=0}^{\infty} \frac{1}{2^i} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2.$$

Other Summations

Other Summations

- Sum of first n integers

Other Summations

- Sum of first n integers

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} = \Theta\left(n^2\right)$$

Other Summations

- Sum of first n integers

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} = \Theta\left(n^2\right)$$

- Sum of first n squares

Other Summations

- Sum of first n integers

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} = \Theta\left(n^2\right)$$

- Sum of first n squares

$$\sum_{i=1}^{n} i^2 = 1 + 4 + 9 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} = \Theta\left(n^3\right)$$

Other Summations

- Sum of first n integers

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} = \Theta\left(n^2\right)$$

- Sum of first n squares

$$\sum_{i=1}^{n} i^2 = 1 + 4 + 9 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} = \Theta\left(n^3\right)$$

- In general, for any fixed positive integer k:

Other Summations

- Sum of first n integers

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} = \Theta\left(n^2\right)$$

- Sum of first n squares

$$\sum_{i=1}^{n} i^2 = 1 + 4 + 9 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} = \Theta\left(n^3\right)$$

- In general, for any fixed positive integer k:

$$\sum_{i=1}^{n} i^k = 1 + 2^k + 3^k + \cdots + n^k = \Theta\left(n^{k+1}\right)$$

Logarithms

Logarithms

Definition: $\log_b x = y$ if and only if $b^y = x$.

Logarithms

Definition: $\log_b x = y$ if and only if $b^y = x$.

Some useful properties:

Logarithms

Definition: $\log_b x = y$ if and only if $b^y = x$.

Some useful properties:

1. $\log_b 1 = 0$.
2. $\log_b b^a = a$.
3. $\log_b(xy) = \log_b x + \log_b y$.
4. $\log_b(x^a) = a \log_b x$.
5. $x^{\log_b y} = y^{\log_b x}$.
6. $\log_x b = \frac{1}{\log_b x}$.
7. $\log_a x = \frac{\log_b x}{\log_b a}$.
8. $\log_a x = (\log_b x)(\log_a b)$.

Logarithms

Definition: $\log_b x = y$ if and only if $b^y = x$.

Some useful properties:

1. $\log_b 1 = 0$.
2. $\log_b b^a = a$.
3. $\log_b(xy) = \log_b x + \log_b y$.
4. $\log_b(x^a) = a \log_b x$.
5. $x^{\log_b y} = y^{\log_b x}$.
6. $\log_x b = \frac{1}{\log_b x}$.
7. $\log_a x = \frac{\log_b x}{\log_b a}$.
8. $\log_a x = (\log_b x)(\log_a b)$.

Exercise: Prove the above properties.

Logarithms

Logarithms

Example (#2): Prove $\log_b b^a = a$.

Logarithms

Example (#2): Prove $\log_b b^a = a$.

Solution: Let $y = \log_b b^a$

Logarithms

Example (#2): Prove $\log_b b^a = a$.

Solution: Let

$$\begin{aligned} y &= \log_b b^a \\ b^y &= b^a \quad \text{[by definition of log]} \end{aligned}$$

Logarithms

Example (#2): Prove $\log_b b^a = a$.

Solution: Let

$$\begin{aligned} y &= \log_b b^a \\ b^y &= b^a \quad \text{[by definition of log]} \\ y &= a \end{aligned}$$

Logarithms

Special Notations:

Logarithms

Special Notations:

- $\ln x = \log_e x$ $(e = 2.71828\ldots)$

Logarithms

Special Notations:

- $\ln x = \log_e x$ ($e = 2.71828\ldots$)
- $\lg x = \log_2 x$

Logarithms

Special Notations:

- $\ln x = \log_e x$ ($e = 2.71828\ldots$)
- $\lg x = \log_2 x$

Some conversions (from Rules #7 and #8 on previous slides):

Logarithms

Special Notations:

- $\ln x = \log_e x$ ($e = 2.71828\ldots$)
- $\lg x = \log_2 x$

Some conversions (from Rules #7 and #8 on previous slides):

- $\ln x = (\log_2 x)(\log_e 2) = 0.693 \lg x$.

Logarithms

Special Notations:

- $\ln x = \log_e x$ ($e = 2.71828\ldots$)
- $\lg x = \log_2 x$

Some conversions (from Rules #7 and #8 on previous slides):

- $\ln x = (\log_2 x)(\log_e 2) = 0.693 \lg x$.
- $\lg x = \frac{\log_e x}{\log_e 2} = \frac{\ln x}{0.693} = 1.44 \ln x$.

Floors and ceilings

Floors and ceilings

- $\lfloor x \rfloor$ = largest integer $\leq x$. (Read as Floor of x)

Floors and ceilings

- $\lfloor x \rfloor$ = largest integer $\leq x$. (Read as Floor of x)
- $\lceil x \rceil$= smallest integer $\geq x$ (Read as Ceiling of x)

Factorials

- $n! = 1 \cdot 2 \cdots n$

Factorials

- $n! = 1 \cdot 2 \cdots n$
- $n!$ represents the number of distinct permutations of n objects.

Factorials

- $n! = 1 \cdot 2 \cdots n$
- $n!$ represents the number of distinct permutations of n objects.

1	2	3
1	3	2
2	1	3
2	3	1
3	1	2
3	2	1

Combinations

Combinations

$\binom{n}{k}$ = The number of different ways of choosing k objects from a collection of n objects. (Pronounced "n choose k".)

Combinations

$\binom{n}{k}$ = The number of different ways of choosing k objects from a collection of n objects. (Pronounced "n choose k".)

Example: $\binom{5}{2} = 10$

Combinations

$\binom{n}{k}$ = The number of different ways of choosing k objects from a collection of n objects. (Pronounced "n choose k".)

Example: $\binom{5}{2} = 10$

$$\begin{array}{ccccc} \{1,2\} & \{1,3\} & \{1,4\} & \{1,5\} & \{2,3\} \\ \{2,4\} & \{2,5\} & \{3,4\} & \{3,5\} & \{4,5\} \end{array}$$

Combinations

$\binom{n}{k}$ = The number of different ways of choosing k objects from a collection of n objects. (Pronounced "n choose k".)

Example: $\binom{5}{2} = 10$

$$\begin{array}{ccccc} \{1,2\} & \{1,3\} & \{1,4\} & \{1,5\} & \{2,3\} \\ \{2,4\} & \{2,5\} & \{3,4\} & \{3,5\} & \{4,5\} \end{array}$$

Formula: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Combinations

$\binom{n}{k}$ = The number of different ways of choosing k objects from a collection of n objects. (Pronounced "n choose k".)

Example: $\binom{5}{2} = 10$

$$\begin{array}{ccccc} \{1,2\} & \{1,3\} & \{1,4\} & \{1,5\} & \{2,3\} \\ \{2,4\} & \{2,5\} & \{3,4\} & \{3,5\} & \{4,5\} \end{array}$$

Formula: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Special cases: $\binom{n}{0} = 1$, $\binom{n}{1} = n$, $\binom{n}{2} = \frac{n(n-1)}{2}$

Harmonic Numbers

Harmonic Numbers

The nth Harmonic number is the sum:

$$H_n = \sum_{i=1}^{n} \frac{1}{i}$$

Harmonic Numbers

The *n*th Harmonic number is the sum:

$$H_n = \sum_{i=1}^{n} \frac{1}{i}$$

These numbers go to infinity:

$$\lim_{n \to \infty} H_n = \sum_{i=1}^{\infty} \frac{1}{i} = \infty$$

Harmonic Numbers

The harmonic numbers are closely related to logs.

Harmonic Numbers

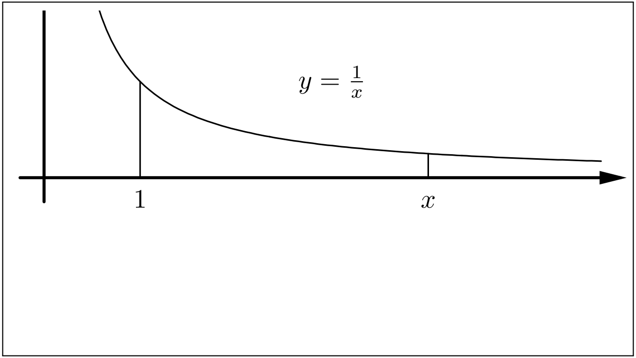
The harmonic numbers are closely related to logs. Recall:

$$\ln x = \int_1^x \frac{1}{t}\,dt$$

Harmonic Numbers

The harmonic numbers are closely related to logs. Recall:

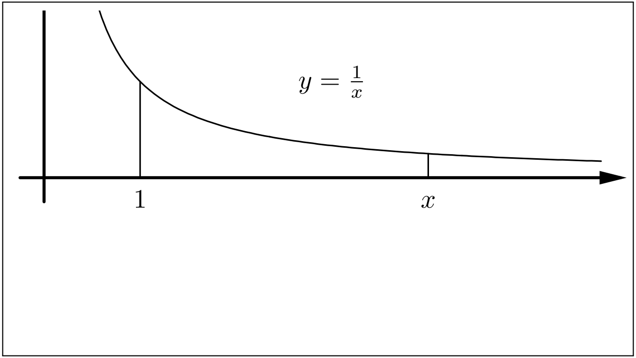
$$\ln x = \int_1^x \frac{1}{t}\,dt$$



Harmonic Numbers

The harmonic numbers are closely related to logs. Recall:

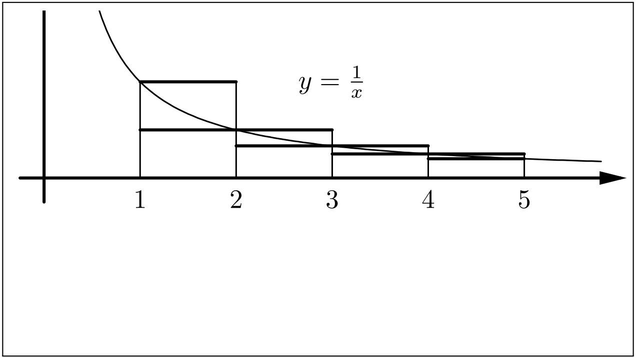
$$\ln x = \int_1^x \frac{1}{t}\,dt$$



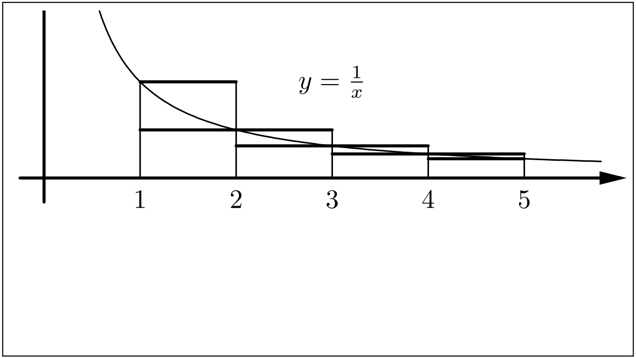
We will show that $H_n = \Theta(\log n)$.

Harmonic Numbers

Harmonic Numbers

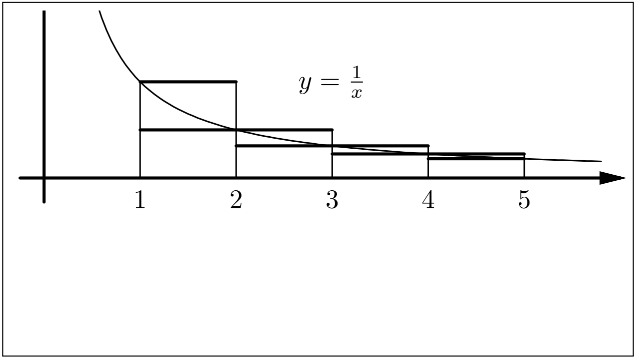


Harmonic Numbers



$$\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} \quad < \quad \ln n \quad < \quad 1 + \frac{1}{2} + \ldots + \frac{1}{n-1}$$

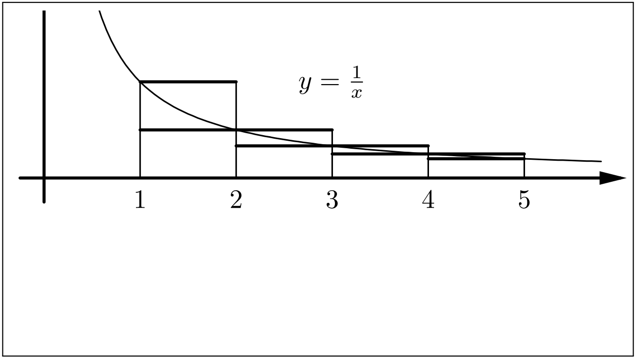
Harmonic Numbers



$$\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} \quad < \quad \ln n \quad < \quad 1 + \frac{1}{2} + \ldots + \frac{1}{n-1}$$

$$H_n - 1 \quad < \quad \ln n \quad < \quad H_n - \frac{1}{n}$$

Harmonic Numbers



$$\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} \quad < \quad \ln n \quad < \quad 1 + \frac{1}{2} + \ldots + \frac{1}{n-1}$$

$$H_n - 1 \quad < \quad \ln n \quad < \quad H_n - \frac{1}{n}$$

Hence $\ln n + \frac{1}{n} < H_n < \ln n + 1$, so $H_n = \Theta(\log n)$.

Proof/Justification Techniques

Proof/Justification Techniques

- Proof by Example Can be used to prove

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.
  - A statement of the form "For all..." is false.

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.
  - A statement of the form "For all..." is false.
  - A statement of the form "If P then Q" is false.

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.
  - A statement of the form "For all..." is false.
  - A statement of the form "If P then Q" is false.
- Illustration:

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.
  - A statement of the form "For all..." is false.
  - A statement of the form "If P then Q" is false.
- Illustration: Consider the statement:

  All numbers of the form $2^k - 1$ are prime.

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.
  - A statement of the form "For all..." is false.
  - A statement of the form "If P then Q" is false.
- Illustration: Consider the statement:

  All numbers of the form $2^k - 1$ are prime.

  This statement is False: $2^4 - 1 = 15 = 3 \cdot 5$

Proof/Justification Techniques

- Proof by Example Can be used to prove
  - A statement of the form "There exists..." is true.
  - A statement of the form "For all..." is false.
  - A statement of the form "If P then Q" is false.
- Illustration: Consider the statement:

  All numbers of the form $2^k - 1$ are prime.

  This statement is False: $2^4 - 1 = 15 = 3 \cdot 5$
- Note: The statement can be rewritten as:

  If n is an integer of the form $2^k - 1$, then n is prime.

Proof/Justification Techniques

- Suppose we want to prove a statement of the form "If P then Q" is true.

Proof/Justification Techniques

- Suppose we want to prove a statement of the form "If P then Q" is true.
  There are three approaches:

Proof/Justification Techniques

- Suppose we want to prove a statement of the form "If P then Q" is true.
  There are three approaches:
  1. Direct proof: Assume P is true. Show that Q must be true.

Proof/Justification Techniques

- Suppose we want to prove a statement of the form "If P then Q" is true.
  There are three approaches:
  1. Direct proof: Assume P is true. Show that Q must be true.
  2. Indirect proof: Assume Q is false. Show that P must be false.

Proof/Justification Techniques

- Suppose we want to prove a statement of the form "If P then Q" is true.
  There are three approaches:
  1. Direct proof: Assume P is true. Show that Q must be true.
  2. Indirect proof: Assume Q is false. Show that P must be false. This is also known as a proof by contraposition.
  3. Proof by contradiction: Assume P is true and Q is false. Show that there is a contradiction.

Proof/Justification Techniques

- Suppose we want to prove a statement of the form "If P then Q" is true.
  There are three approaches:
  1. Direct proof: Assume P is true. Show that Q must be true.
  2. Indirect proof: Assume Q is false. Show that P must be false. This is also known as a proof by contraposition.
  3. Proof by contradiction: Assume P is true and Q is false. Show that there is a contradiction.

  See [GT] Section 1.3.3 for examples.

Proof/Justification Techniques: Induction

Proof/Justification Techniques: Induction

- A technique for proving theorems about the positive (or nonnegative) integers.

Proof/Justification Techniques: Induction

- A technique for proving theorems about the positive (or nonnegative) integers.
- Let $P(n)$ be a statement with an integer parameter, n. Mathematical induction is a technique for proving that $P(n)$ is true for all integers $\geq$ some base value b.

Proof/Justification Techniques: Induction

- A technique for proving theorems about the positive (or nonnegative) integers.
- Let $P(n)$ be a statement with an integer parameter, n. Mathematical induction is a technique for proving that $P(n)$ is true for all integers $\geq$ some base value b.
- Usually, the base value is 0 or 1.

Proof/Justification Techniques: Induction

- A technique for proving theorems about the positive (or nonnegative) integers.
- Let $P(n)$ be a statement with an integer parameter, n. Mathematical induction is a technique for proving that $P(n)$ is true for all integers $\geq$ some base value b.
- Usually, the base value is 0 or 1.
- To show $P(n)$ holds for all $n \geq b$, we must show two things:

Proof/Justification Techniques: Induction

- A technique for proving theorems about the positive (or nonnegative) integers.
- Let $P(n)$ be a statement with an integer parameter, n. Mathematical induction is a technique for proving that $P(n)$ is true for all integers $\geq$ some base value b.
- Usually, the base value is 0 or 1.
- To show $P(n)$ holds for all $n \geq b$, we must show two things:
  1. Base Case: $P(b)$ is true (where b is the base value).

Proof/Justification Techniques: Induction

- A technique for proving theorems about the positive (or nonnegative) integers.
- Let $P(n)$ be a statement with an integer parameter, n. Mathematical induction is a technique for proving that $P(n)$ is true for all integers $\geq$ some base value b.
- Usually, the base value is 0 or 1.
- To show $P(n)$ holds for all $n \geq b$, we must show two things:
    1. Base Case: $P(b)$ is true (where b is the base value).
    2. Inductive step: If $P(k)$ is true, then $P(k+1)$ is true.

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

LHS

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

$$\text{LHS} \;=\; \sum_{i=1}^{1} i \cdot 2^i = 1 \cdot 2^1 = 2.$$

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

$$\text{LHS} = \sum_{i=1}^{1} i \cdot 2^i = 1 \cdot 2^1 = 2.$$

RHS

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

$$\begin{aligned} \text{LHS} &= \sum_{i=1}^{1} i \cdot 2^i = 1 \cdot 2^1 = 2. \\ \text{RHS} &= (1-1) \cdot 2^{1+1} + 2 = 0 + 2 = 2. \end{aligned}$$

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

$$\begin{aligned}
\text{LHS} &= \sum_{i=1}^{1} i \cdot 2^i = 1 \cdot 2^1 = 2. \\
\text{RHS} &= (1-1) \cdot 2^{1+1} + 2 = 0 + 2 = 2. \\
\text{LHS} &
\end{aligned}$$

Induction Example

Example: Show that for all $n \geq 1$

$$\sum_{i=1}^{n} i \cdot 2^i = (n-1) \cdot 2^{(n+1)} + 2$$

Base Case: $(n = 1)$

$$\begin{aligned} \text{LHS} &= \sum_{i=1}^{1} i \cdot 2^i = 1 \cdot 2^1 = 2. \\ \text{RHS} &= (1-1) \cdot 2^{1+1} + 2 = 0 + 2 = 2. \\ \text{LHS} &= \text{RHS} \quad \checkmark \end{aligned}$$

Induction Example, continued

Inductive Step:

Induction Example, continued

Inductive Step:

Assume $P(k)$ is true:

$$\sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

Induction Example, continued

Inductive Step:

Assume $P(k)$ is true:

$$\sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

Show $P(k+1)$ is true:

$$\sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

Induction Example, continued

$$\text{Assume: } \sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

$$\text{Show: } \quad \sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

Induction Example, continued

$$\text{Assume: } \sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

$$\text{Show: } \quad \sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

$$\sum_{i=1}^{k+1} i \cdot 2^i$$

Induction Example, continued

$$\text{Assume: } \sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

$$\text{Show: } \quad \sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

$$\sum_{i=1}^{k+1} i \cdot 2^i \quad = \quad \sum_{i=1}^{k} i \cdot 2^i + (k+1) \cdot 2^{(k+1)}$$

Induction Example, continued

$$\text{Assume: } \sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

$$\text{Show: } \quad \sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

$$\begin{aligned} \sum_{i=1}^{k+1} i \cdot 2^i &= \sum_{i=1}^{k} i \cdot 2^i + (k+1) \cdot 2^{(k+1)} \\ &= (k-1) \cdot 2^{(k+1)} + 2 + (k+1) \cdot 2^{(k+1)} \end{aligned}$$

Induction Example, continued

$$\text{Assume: } \sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

$$\text{Show: } \quad \sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

$$\begin{aligned}
\sum_{i=1}^{k+1} i \cdot 2^i &= \sum_{i=1}^{k} i \cdot 2^i + (k+1) \cdot 2^{(k+1)} \\
&= (k-1) \cdot 2^{(k+1)} + 2 + (k+1) \cdot 2^{(k+1)} \\
&= 2k \cdot 2^{(k+1)} + 2
\end{aligned}$$

Induction Example, continued

$$\text{Assume: } \sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$$

$$\text{Show: } \sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$$

$$\begin{aligned}
\sum_{i=1}^{k+1} i \cdot 2^i &= \sum_{i=1}^{k} i \cdot 2^i + (k+1) \cdot 2^{(k+1)} \\
&= (k-1) \cdot 2^{(k+1)} + 2 + (k+1) \cdot 2^{(k+1)} \\
&= 2k \cdot 2^{(k+1)} + 2 \\
&= k \cdot 2^{(k+2)} + 2
\end{aligned}$$

Induction Example, continued

Assume: $\sum_{i=1}^{k} i \cdot 2^i = (k-1) \cdot 2^{(k+1)} + 2.$

Show: $\sum_{i=1}^{k+1} i \cdot 2^i = k \cdot 2^{(k+2)} + 2.$

$$
\begin{aligned}
\sum_{i=1}^{k+1} i \cdot 2^i &= \sum_{i=1}^{k} i \cdot 2^i + (k+1) \cdot 2^{(k+1)} \\
&= (k-1) \cdot 2^{(k+1)} + 2 + (k+1) \cdot 2^{(k+1)} \\
&= 2k \cdot 2^{(k+1)} + 2 \\
&= k \cdot 2^{(k+2)} + 2 \quad \text{QED}
\end{aligned}
$$

Probability

Probability

- Defined in terms of a sample space, S.

Probability

- Defined in terms of a sample space, S.
- Sample space consists of a finite set of outcomes, also called elementary events.

Probability

- Defined in terms of a sample space, S.
- Sample space consists of a finite set of outcomes, also called elementary events.
- An event is a subset of the sample space. (So an event is a set of outcomes).

Probability

- Defined in terms of a sample space, S.
- Sample space consists of a finite set of outcomes, also called elementary events.
- An event is a subset of the sample space. (So an event is a set of outcomes).
- Sample space can be infinite, even uncountable. In this course, it will generally be finite.

Probability

- Defined in terms of a sample space, S.
- Sample space consists of a finite set of outcomes, also called elementary events.
- An event is a subset of the sample space. (So an event is a set of outcomes).
- Sample space can be infinite, even uncountable. In this course, it will generally be finite.

Example: (2-coin example.) Flip two coins.

Probability

- Defined in terms of a sample space, S.
- Sample space consists of a finite set of outcomes, also called elementary events.
- An event is a subset of the sample space. (So an event is a set of outcomes).
- Sample space can be infinite, even uncountable. In this course, it will generally be finite.

Example: (2-coin example.) Flip two coins.

- Sample space $S = \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}$.

Probability

- Defined in terms of a sample space, S.
- Sample space consists of a finite set of outcomes, also called elementary events.
- An event is a subset of the sample space. (So an event is a set of outcomes).
- Sample space can be infinite, even uncountable. In this course, it will generally be finite.

Example: (2-coin example.) Flip two coins.

- Sample space $S = \{\mathtt{HH}, \mathtt{HT}, \mathtt{TH}, \mathtt{TT}\}$.
- The event "first coin is heads" is the subset $\{\mathtt{HH}, \mathtt{HT}\}$.

Probability function

Probability function

- A probability function is a function $P(\cdot)$ that maps events (subsets of the sample space S) to real numbers such that:

Probability function

- A probability function is a function $P(\cdot)$ that maps events (subsets of the sample space S) to real numbers such that:
  1. $P(\emptyset) = 0$.

Probability function

- A probability function is a function $P(\cdot)$ that maps events (subsets of the sample space S) to real numbers such that:
  1. $P(\emptyset) = 0$.
  2. $P(S) = 1$.

Probability function

- A probability function is a function $P(\cdot)$ that maps events (subsets of the sample space S) to real numbers such that:
  1. $P(\emptyset) = 0$.
  2. $P(S) = 1$.
  3. For every event A, $0 \leq P(A) \leq 1$.

Probability function

- A probability function is a function $P(\cdot)$ that maps events (subsets of the sample space S) to real numbers such that:
  1. $P(\emptyset) = 0$.
  2. $P(S) = 1$.
  3. For every event A, $0 \leq P(A) \leq 1$.
  4. If $A, B \subseteq S$ and $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

Probability function

- A probability function is a function $P(\cdot)$ that maps events (subsets of the sample space S) to real numbers such that:
  1. $P(\emptyset) = 0$.
  2. $P(S) = 1$.
  3. For every event A, $0 \leq P(A) \leq 1$.
  4. If $A, B \subseteq S$ and $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.
- Note: Property 4 implies that if $A \subseteq B$ then $P(A) \leq P(B)$.

Probability function (continued)

Probability function (continued)

For finite sample spaces, this can be simplified:

Probability function (continued)

For finite sample spaces, this can be simplified:

- Sample space $S = \{s_1, \ldots, s_k\}$,

Probability function (continued)

For finite sample spaces, this can be simplified:

- Sample space $S = \{s_1, \ldots, s_k\}$,
- Each outcome S_i is assigned a probability $P(s_i)$, with

$$\sum_{i=1}^{k} P(s_i) = 1.$$

Probability function (continued)

For finite sample spaces, this can be simplified:

- Sample space $S = \{s_1, \ldots, s_k\}$,
- Each outcome S_i is assigned a probability $P(s_i)$, with

$$\sum_{i=1}^{k} P(s_i) = 1.$$

- The probability of an event $E \subseteq S$ is:

$$P(E) = \sum_{s_i \in E} P(s_i).$$

Probability function (continued)

For finite sample spaces, this can be simplified:

- Sample space $S = \{s_1, \ldots, s_k\}$,
- Each outcome S_i is assigned a probability $P(s_i)$, with

$$\sum_{i=1}^{k} P(s_i) = 1.$$

- The probability of an event $E \subseteq S$ is:

$$P(E) = \sum_{s_i \in E} P(s_i).$$

Example: (2-coin example, continued). Define

$$P(\mathtt{HH}) = P(\mathtt{HT}) = P(\mathtt{TH}) = P(\mathtt{TT}) = \frac{1}{4}.$$

Probability function (continued)

For finite sample spaces, this can be simplified:

- Sample space $S = \{s_1, \ldots, s_k\}$,
- Each outcome S_i is assigned a probability $P(s_i)$, with

$$\sum_{i=1}^{k} P(s_i) = 1.$$

- The probability of an event $E \subseteq S$ is:

$$P(E) = \sum_{s_i \in E} P(s_i).$$

Example: (2-coin example, continued). Define

$$P(\mathtt{HH}) = P(\mathtt{HT}) = P(\mathtt{TH}) = P(\mathtt{TT}) = \frac{1}{4}.$$

Then

$$P(\text{first coin is heads}) = P(\mathtt{HH}) + P(\mathtt{HT}) = \frac{1}{2}.$$

Random variables

Random variables

- Intuitive definition: a random variable is a variable whose value depends on the outcome of some experiment.

Random variables

- Intuitive definition: a random variable is a variable whose value depends on the outcome of some experiment.
- Formal definition: a random variable is a function that maps outcomes in a sample space S to real numbers.

Random variables

- Intuitive definition: a random variable is a variable whose value depends on the outcome of some experiment.
- Formal definition: a random variable is a function that maps outcomes in a sample space S to real numbers.
- Special case: An Indicator variable is a random variable that is always either 0 or 1.

Expectation

Expectation

- The expected value, or expectation, of a random variable X represents its "average value".

Expectation

- The expected value, or expectation, of a random variable X represents its "average value".
- Formally: Let X be a random variable with a finite set of possible values $\mathrm{V} = \{x_1, \ldots, x_k\}$. Then

$$E(X) = \sum_{x \in V} x \cdot P(X = x).$$

Expectation

- The expected value, or expectation, of a random variable X represents its "average value".
- Formally: Let X be a random variable with a finite set of possible values $\mathrm{V} = \{x_1, \ldots, x_k\}$. Then

$$E(X) = \sum_{x \in V} x \cdot P(X = x).$$

Example: (2-coin example, continued). Let X be the number of heads when two coins are thrown. Then

$$\begin{aligned} E(X) &= 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) \\ &= 0 \cdot \left(\frac{1}{4}\right) + 1 \cdot \left(\frac{1}{2}\right) + 2 \cdot \left(\frac{1}{4}\right) \\ &= 1 \end{aligned}$$

Expectation

Expectation

Example: Throw a single six-sided die. Assume the die is fair, so each possible throw has a probability of $1/6$.

Expectation

Example: Throw a single six-sided die. Assume the die is fair, so each possible throw has a probability of $1/6$.

The expected value of the throw is:

$$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

Linearity of Expectation

Linearity of Expectation

- For any two random variables X and Y,

$$E(X+Y) = E(X) + E(Y).$$

Linearity of Expectation

- For any two random variables X and Y,

$$E(X + Y) = E(X) + E(Y).$$

- Proof: see [GT], 1.3.4

Linearity of Expectation

- For any two random variables X and Y,

$$E(X+Y) = E(X) + E(Y).$$

- Proof: see [GT], 1.3.4
- Very useful, because usually it is easier to compute $E(X)$ and $E(Y)$ and apply the formula than to compute $E(X+Y)$ directly.

Linearity of Expectation

- For any two random variables X and Y,

$$E(X+Y) = E(X) + E(Y).$$

- Proof: see [GT], 1.3.4
- Very useful, because usually it is easier to compute $E(X)$ and $E(Y)$ and apply the formula than to compute $E(X+Y)$ directly.

Example 1: Throw two six-sided dice. Let X be the sum of the values. Then

$$E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = 3.5 + 3.5 = 7,$$

where X_i is the value on die i ($i = 1, 2$).

Linearity of Expectation

- For any two random variables X and Y,

$$E(X+Y) = E(X) + E(Y).$$

- Proof: see [GT], 1.3.4
- Very useful, because usually it is easier to compute $E(X)$ and $E(Y)$ and apply the formula than to compute $E(X+Y)$ directly.

Example 1: Throw two six-sided dice. Let X be the sum of the values. Then

$$E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = 3.5 + 3.5 = 7,$$

where X_i is the value on die i ($i = 1, 2$).

Example 2: Throw 100 six-sided dice. Let Y be the sum of the values. Then

$$E(Y) = 100 \cdot 3.5 = 350.$$

Independent events

Independent events

- Two events A_1 and A_2 are independent iff

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2).$$

Example: (2-coin example, continued).

Independent events

- Two events A_1 and A_2 are independent iff

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2).$$

Example: (2-coin example, continued). Let

$$\begin{array}{lclcl} A_1 & = & \text{coin 1 is heads} & = & \{\mathtt{HH}, \mathtt{HT}\} \\ A_2 & = & \text{coin 2 is tails} & = & \{\mathtt{HT}, \mathtt{TT}\} \end{array}$$

Independent events

- Two events A_1 and A_2 are independent iff

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2).$$

Example: (2-coin example, continued). Let

$$\begin{array}{lclcl} A_1 & = & \text{coin 1 is heads} & = & \{\texttt{HH}, \texttt{HT}\} \\ A_2 & = & \text{coin 2 is tails} & = & \{\texttt{HT}, \texttt{TT}\} \end{array}$$

Then $P(A_1) = \frac{1}{2}$, $P(A_2) = \frac{1}{2}$, and

$$P(A_1 \cap A_2) = P(\texttt{HT}) = \frac{1}{4} = P(A_1) \cdot P(A_2).$$

So A_1 and A_2 are independent.

Independent events

Independent events

A collection of n events $C = \{A_1, A_2, \ldots, A_n\}$ is mutually independent (or simply independent) if:

For every subset $\{A_{i_1}, A_{i_2}, \ldots A_{i_k}\} \subseteq C$*:*

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k}).$$

Independent events

A collection of n events $C = \{A_1, A_2, \ldots, A_n\}$ is mutually independent (or simply independent) if:

For every subset $\{A_{i_1}, A_{i_2}, \ldots A_{i_k}\} \subseteq C$*:*

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k}).$$

Example: Suppose we flip 10 coins. Suppose the flips are fair ($P(\mathtt{H}) = P(\mathtt{T}) = 1/2$) and independent. Then the probability of any particular sequence of flips (e.g., `HHTTTHTHTH`) is $1/(2^{10})$.

Example: Probability and counting

Example: Probability and counting

Example: Suppose we flip a coin 10 times. Suppose the flips are fair and independent. What is the probability of getting exactly 7 heads out of the 10 flips?

Example: Probability and counting

Example: Suppose we flip a coin 10 times. Suppose the flips are fair and independent. What is the probability of getting exactly 7 heads out of the 10 flips?

Solution:

- The outcomes consist of the set of possible sequences of 10 flips (e.g., `HHTTTHTHTH`).

Example: Probability and counting

Example: Suppose we flip a coin 10 times. Suppose the flips are fair and independent. What is the probability of getting exactly 7 heads out of the 10 flips?

Solution:

- The outcomes consist of the set of possible sequences of 10 flips (e.g., `HHTTTHTHTH`).
- The probability of each outcome is $1/(2^{10})$.

Example: Probability and counting

Example: Suppose we flip a coin 10 times. Suppose the flips are fair and independent. What is the probability of getting exactly 7 heads out of the 10 flips?

Solution:

- The outcomes consist of the set of possible sequences of 10 flips (e.g., HHTTTHTHTH).
- The probability of each outcome is $1/(2^{10})$.
- The number of successful outcomes is $\binom{10}{7}$.

Example: Probability and counting

Example: Suppose we flip a coin 10 times. Suppose the flips are fair and independent. What is the probability of getting exactly 7 heads out of the 10 flips?

Solution:

- The outcomes consist of the set of possible sequences of 10 flips (e.g., `HHTTTHTHTH`).
- The probability of each outcome is $1/(2^{10})$.
- The number of successful outcomes is $\binom{10}{7}$.
- Hence the probability of getting exactly 7 heads is:

$$\frac{\binom{10}{7}}{2^{10}} = \frac{120}{1024} = 0.117.$$

An average-case result about finding the maximum

An average-case result about finding the maximum

```
v = -∞
for i = 0 to n-1:
   if A[i] > v:
      v = A[i]
return v
```

An average-case result about finding the maximum

```
v = -∞
for i = 0 to n-1:
   if A[i] > v:
      v = A[i]
return v
```

- Worst-case number of comparisons is n.

An average-case result about finding the maximum

```
v = -∞
for i = 0 to n-1:
   if A[i] > v:
      v = A[i]
return v
```

- Worst-case number of comparisons is n.
- This can be reduced to $n-1$

An average-case result about finding the maximum

```
v = -∞
for i = 0 to n-1:
   if A[i] > v:
      v = A[i]
return v
```

- Worst-case number of comparisons is n.
- This can be reduced to $n-1$
- How many times is the running maximum updated?

An average-case result about finding the maximum

```
v = -∞
for i = 0 to n-1:
    if A[i] > v:
        v = A[i]
return v
```

- Worst-case number of comparisons is n.
- This can be reduced to $n-1$
- How many times is the running maximum updated?
  - In the worst case n.

An average-case result about finding the maximum

```
v = -∞
for i = 0 to n-1:
   if A[i] > v:
      v = A[i]
return v
```

- Worst-case number of comparisons is n.
- This can be reduced to $n-1$
- How many times is the running maximum updated?
  - In the worst case n.
  - What about the average case? ...

Average number of updates to the running maximum

Average number of updates to the running maximum

- Assume
  - all possible orderings (permutations) of A are equally likely
  - all n elements of A are distinct.

Average number of updates to the running maximum

- Assume
  - all possible orderings (permutations) of A are equally likely
  - all n elements of A are distinct.
- The running maximum gets updated on iteration i of the loop iff $\max\{A[0], \ldots, A[i]\} = A[i]$.

Average number of updates to the running maximum

- Assume
  - all possible orderings (permutations) of A are equally likely
  - all n elements of A are distinct.
- The running maximum gets updated on iteration i of the loop iff $\max\{A[0], \ldots, A[i]\} = A[i]$.
- The probability of this happening is $1/(i+1)$.

Average number of updates to the running maximum

- Assume
  - all possible orderings (permutations) of A are equally likely
  - all n elements of A are distinct.
- The running maximum gets updated on iteration i of the loop iff $\max\{A[0], \ldots, A[i]\} = A[i]$.
- The probability of this happening is $1/(i+1)$.
- Define indicator variables X_i:

$$X_i = \begin{cases} 1 & \text{if } v \text{ gets updated on iteration } \#i \\ 0 & \text{if } v \text{ does not get updated on iteration } \#i \end{cases}$$

Average number of updates to the running maximum

- Assume
  - all possible orderings (permutations) of A are equally likely
  - all n elements of A are distinct.
- The running maximum gets updated on iteration i of the loop iff $\max\{A[0], \ldots, A[i]\} = A[i]$.
- The probability of this happening is $1/(i+1)$.
- Define indicator variables X_i:

$$X_i = \begin{cases} 1 & \text{if } v \text{ gets updated on iteration } \#i \\ 0 & \text{if } v \text{ does not get updated on iteration } \#i \end{cases}$$

Then $E(X_i) = \frac{1}{i+1}$

Average number of updates to the running maximum

- Assume
  - all possible orderings (permutations) of A are equally likely
  - all n elements of A are distinct.
- The running maximum gets updated on iteration i of the loop iff $\max\{A[0], \ldots, A[i]\} = A[i]$.
- The probability of this happening is $1/(i+1)$.
- Define indicator variables X_i:

$$X_i = \begin{cases} 1 & \text{if } v \text{ gets updated on iteration } \#i \\ 0 & \text{if } v \text{ does not get updated on iteration } \#i \end{cases}$$

  Then $E(X_i) = \frac{1}{i+1}$

- The total number of times that v gets updated is:

$$X = \sum_{i=0}^{n-1} X_i$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$E(X)$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right)$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right)$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1}$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1} = \sum_{i=1}^{n} \frac{1}{i}$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1} = \sum_{i=1}^{n} \frac{1}{i} = H_n$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1} = \sum_{i=1}^{n} \frac{1}{i} = H_n = O(\log n)$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1} = \sum_{i=1}^{n} \frac{1}{i} = H_n = O(\log n)$$

It can be shown that

$$H_n = \ln n + \gamma + o(1), \quad \text{where } \gamma = 0.5772157\ldots$$

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1} = \sum_{i=1}^{n} \frac{1}{i} = H_n = O(\log n)$$

It can be shown that

$$H_n = \ln n + \gamma + o(1), \quad \text{where } \gamma = 0.5772157\ldots$$

If there are 30,000 elements in the list, the expected update count is about 10.9

Average number of updates to the running maximum (continued)

The expected total number of times that v gets updated is:

$$E(X) = E\left(\sum_{i=0}^{n-1} X_i\right) = \sum_{i=0}^{n-1} E\left(X_i\right) = \sum_{i=0}^{n-1} \frac{1}{i+1} = \sum_{i=1}^{n} \frac{1}{i} = H_n = O(\log n)$$

It can be shown that

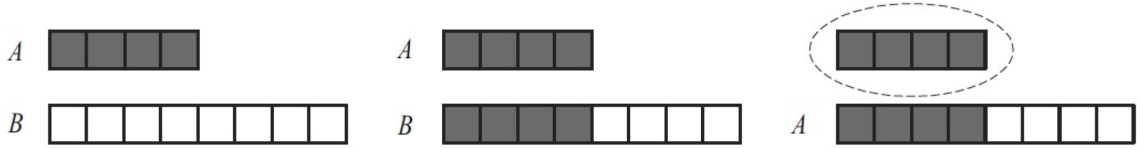
$$H_n = \ln n + \gamma + o(1), \quad \text{where } \gamma = 0.5772157\ldots$$

If there are 30,000 elements in the list, the expected update count is about 10.9

If there are 3,000,000,000 elements in the list, the expected update count is about 22.4

Amortization

- The **amortized running time** of an operation within a series of operations is the worst-case running time of the series of operations divided by the number of operations.
- Example: A growable array, S. When needing to grow:
  a. Allocate a new array B of larger capacity.
  b. Copy A[i] to B[i], for i = 0, . . . , n − 1, where n is size of A.
  c. Let A = B, that is, we use B as the array now supporting A.



(a) (b) (c)

Growable Array Description

- Let add(e) be the operation that adds element e at the end of the array
- When the array is full, we replace the array with a larger one
- But how large should the new array be?
  - Incremental strategy: increase the size by a constant c
  - Doubling strategy: double the size

Algorithm *add*(*e*)
  if $t = A.length - 1$ **then**
    $B \leftarrow$ **new array of size …**
    for $i \leftarrow 0$ **to** $n-1$ **do**
      $B[i] \leftarrow A[i]$
    $A \leftarrow B$
  $n \leftarrow n + 1$
  $A[n-1] \leftarrow e$

Comparison of the Strategies

- We compare the incremental strategy and the doubling strategy by analyzing the total time $T(n)$ needed to perform a series of n add operations
- We assume that we start with an empty list represented by a growable array of size 1
- We call amortized time of an add operation the average time taken by an add operation over the series of operations, i.e., $T(n)/n$

Incremental Strategy Analysis

- Over n add operations, we replace the array $k = n/c$ times, where c is a constant
- The total time $T(n)$ of a series of n add operations is proportional to

$$n + c + 2c + 3c + 4c + \ldots + kc =$$

$$n + c(1 + 2 + 3 + \ldots + k) =$$

$$n + ck(k + 1)/2$$

- Since c is a constant, $T(n)$ is $O(n + k^2)$, i.e., $O(n^2)$
- Thus, the amortized time of an add operation is $O(n)$

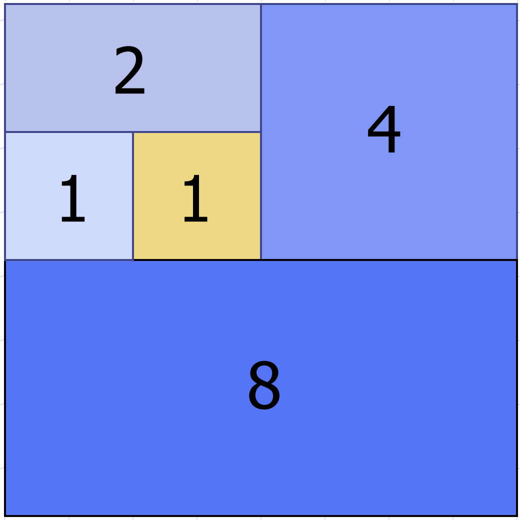
Doubling Strategy Analysis

- We replace the array $k = \log_2 n$ times
- The total time $T(n)$ of a series of n push operations is proportional to

$$n + 1 + 2 + 4 + 8 + \ldots + 2^k =$$
$$n + 2^{k+1} - 1 =$$
$$3n - 1$$

- $T(n)$ is $O(n)$
- The amortized time of an add operation is $O(1)$

geometric series



Accounting Method Proof for the Doubling Strategy

- We view the computer as a coin-operated appliance that requires the payment of 1 **cyber-dollar** for a constant amount of computing time.
- We shall charge each add operation 3 cyber-dollars, that is, it will have an amortized O(1) amortized running time.
  - We over-charge each add operation not causing an overflow 2 cyber-dollars.
  - Think of the 2 cyber-dollars profited in an insertion that does not grow the array as being "stored" at the element inserted.
  - An overflow occurs when the array A has 2^i elements.
  - Thus, doubling the size of the array will require 2^i cyber-dollars.
  - These cyber-dollars are at the elements stored in cells 2^{i-1} through 2^i-1.