Teams

Task: answer questions to get team points
- Score kept during semester
- Periodic immediate rewards

Divide room into 4 teams
- rearrange now if you like,
- sit same place in later classes,
- double agents allowed from my point of view, and
- citizen-exchange allowed from my point of view.
Teams

Choose Team Captain

Choose Team Names

- A: Jason, “The Big O’s”
- B: Andy, “Team Rocket”
- C: Andrew, “Team C”
- D: Victoria, “Algorithm Maniacs”
First Possible Team Point

You have 43 fish in a pond.
17 fish drown.

[TEAMS] How many fish are still alive?
43

(was told to me by my 7-year old son)
Analysis of Algorithms

An algorithm is a step-by-step procedure for solving a problem in a finite amount of time.
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 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
- Use a function, like the built-in `clock()` function, to get an accurate measure of the actual running time
- 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

Example: find max element of an array

```
Algorithm arrayMax(A, n)
Input array A of n integers
Output maximum element of A

currentMax ← A[0]
for i ← 1 to n − 1 do
    if A[i] > currentMax then
        currentMax ← A[i]
return currentMax
```

Example: find max element of an array
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/Function call**
  
  `var.method (arg [, arg...])`

- **Return value**
  
  `return expression`

- **Expressions**
  
  - Assignment (like `=` in C++)
  - Equality testing (like `==` in C++)
  - \( n^2 \) Superscripts and other mathematical formatting allowed

Analysis of Algorithms 10
The Random Access Machine (RAM) Model

- **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.
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

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)`

```plaintext
currentMax ← A[0]
for i ← 1 to n − 1 do
    if A[i] > currentMax then
        currentMax ← A[i]
    
    { increment counter i }

return currentMax
```

<table>
<thead>
<tr>
<th># operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
</tr>
<tr>
<td>2 + n</td>
</tr>
<tr>
<td>2(n − 1)</td>
</tr>
<tr>
<td>2(n − 1)</td>
</tr>
<tr>
<td>2(n − 1)</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Total: \(7n − 1\)
Estimating Running Time

Algorithm *arrayMax* executes $7n - 1$ primitive operations in the worst case.

Define:

- $a = \text{Time taken by the fastest primitive operation}$
- $b = \text{Time taken by the slowest primitive operation}$

Let $T(n)$ be worst-case time of *arrayMax*. Then

$$a \ (7n - 1) \leq T(n) \leq b(7n - 1)$$

Hence, the running time $T(n)$ is bounded by two linear functions.
Power Law

Log-log plot. Plot running time $T(N)$ vs. input size $N$ using log-log scale.
Power Law

- **Log-log plot.** Plot running time $T(N)$ vs. input size $N$ using log-log scale.

- **Regression.** Fit straight line through data points: $a N^b$.

- **Hypothesis.** The running time is about $1.006 \times 10^{-10} \times N^{2.999}$ seconds.

Let $T(N) = a N^b$, where $a = 2^c$

Thus $\log(T(N)) = b \log N + c$

After regression:

$b = 2.999$
$c = -33.2103$
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
Growth Rates

- Growth rates of functions:
  - Linear $\approx n$
  - Quadratic $\approx n^2$
  - Cubic $\approx n^3$

- In a log-log chart, the slope of the line corresponds to the growth rate of the function
Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>problem size solvable in minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1970s</td>
</tr>
<tr>
<td>1</td>
<td>any</td>
</tr>
<tr>
<td>\log N</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>millions</td>
</tr>
<tr>
<td>N \log N</td>
<td>hundreds of thousands</td>
</tr>
<tr>
<td>N^2</td>
<td>hundreds</td>
</tr>
<tr>
<td>N^3</td>
<td>hundred</td>
</tr>
<tr>
<td>2^N</td>
<td></td>
</tr>
</tbody>
</table>
## Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>problem size solvable in minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1970s</td>
</tr>
<tr>
<td>1</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>millions</td>
</tr>
<tr>
<td>N log N</td>
<td>hundreds of thousands</td>
</tr>
<tr>
<td>N^2</td>
<td>hundreds</td>
</tr>
<tr>
<td>N^3</td>
<td>hundred</td>
</tr>
<tr>
<td>2^N</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1980s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>tens of millions</td>
</tr>
<tr>
<td>N log N</td>
<td>millions</td>
</tr>
<tr>
<td>N^2</td>
<td>thousand</td>
</tr>
<tr>
<td>N^3</td>
<td>hundreds</td>
</tr>
<tr>
<td>2^N</td>
<td>20s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1990s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>hundreds of millions</td>
</tr>
<tr>
<td>N log N</td>
<td>millions</td>
</tr>
<tr>
<td>N^2</td>
<td>thousands</td>
</tr>
<tr>
<td>N^3</td>
<td>thousand</td>
</tr>
<tr>
<td>2^N</td>
<td>20s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2000s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>billions</td>
</tr>
<tr>
<td>N log N</td>
<td>hundreds of millions</td>
</tr>
<tr>
<td>N^2</td>
<td>tens of thousands</td>
</tr>
<tr>
<td>N^3</td>
<td>thousands</td>
</tr>
<tr>
<td>2^N</td>
<td>30</td>
</tr>
</tbody>
</table>

Bottom line. Need linear or linearithmic alg to keep pace with Moore’s law.
## Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>name</th>
<th>description</th>
<th>effect on a program that runs for a few seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>time for 100x more data</td>
</tr>
<tr>
<td>1</td>
<td>constant</td>
<td>independent of input size</td>
<td>–</td>
</tr>
<tr>
<td>log N</td>
<td>logarithmic</td>
<td>nearly independent of input size</td>
<td>–</td>
</tr>
<tr>
<td>N</td>
<td>linear</td>
<td>optimal for N inputs</td>
<td>a few minutes</td>
</tr>
<tr>
<td>N log N</td>
<td>linearithmic</td>
<td>nearly optimal for N inputs</td>
<td>a few minutes</td>
</tr>
<tr>
<td>N^2</td>
<td>quadratic</td>
<td>not practical for large problems</td>
<td>several hours</td>
</tr>
<tr>
<td>N^3</td>
<td>cubic</td>
<td>not practical for medium problems</td>
<td>several weeks</td>
</tr>
<tr>
<td>2^N</td>
<td>exponential</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>name</th>
<th>description</th>
<th>effect on a program that runs for a few seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>time for 100x more data</td>
</tr>
<tr>
<td>1</td>
<td>constant</td>
<td>independent of input size</td>
<td>–</td>
</tr>
<tr>
<td>log N</td>
<td>logarithmic</td>
<td>nearly independent of input size</td>
<td>–</td>
</tr>
<tr>
<td>N</td>
<td>linear</td>
<td>optimal for N inputs</td>
<td>a few minutes</td>
</tr>
<tr>
<td>N log N</td>
<td>linearithmic</td>
<td>nearly optimal for N inputs</td>
<td>a few minutes</td>
</tr>
<tr>
<td>N^2</td>
<td>quadratic</td>
<td>not practical for large problems</td>
<td>several hours</td>
</tr>
<tr>
<td>N^3</td>
<td>cubic</td>
<td>not practical for medium problems</td>
<td>several weeks</td>
</tr>
<tr>
<td>2^N</td>
<td>exponential</td>
<td>useful only for tiny problems</td>
<td>forever</td>
</tr>
</tbody>
</table>
## Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>problem size solvable in minutes</th>
<th>time to process millions of inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>any</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>millions</td>
<td>tens of millions</td>
</tr>
<tr>
<td>N log N</td>
<td>hundreds of thousands</td>
<td>millions</td>
</tr>
<tr>
<td>N^2</td>
<td>hundreds</td>
<td>thousand</td>
</tr>
<tr>
<td>N^3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>problem size solvable in minutes</th>
<th>time to process millions of inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>any</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>millions</td>
<td>tens of millions</td>
</tr>
<tr>
<td>N log N</td>
<td>hundreds of thousands</td>
<td>millions</td>
</tr>
<tr>
<td>N^2</td>
<td>hundreds</td>
<td>thousand</td>
</tr>
<tr>
<td>N^3</td>
<td>hundred</td>
<td>hundreds</td>
</tr>
</tbody>
</table>
Practical implications of order-of-growth

<table>
<thead>
<tr>
<th>growth rate</th>
<th>problem size solvable in minutes</th>
<th>time to process millions of inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>any</td>
<td>any</td>
</tr>
<tr>
<td>log N</td>
<td>any</td>
<td>any</td>
</tr>
<tr>
<td>N</td>
<td>millions</td>
<td>tens of millions</td>
</tr>
<tr>
<td>N log N</td>
<td>hundreds of thousands</td>
<td>millions</td>
</tr>
<tr>
<td>N²</td>
<td>hundreds</td>
<td>thousand</td>
</tr>
<tr>
<td>N³</td>
<td>hundred</td>
<td>hundreds</td>
</tr>
</tbody>
</table>
Constant Factors

The growth rate is not affected by
- constant factors or
- lower-order terms

Examples
- $10^2n + 10^5$ is a linear function
- $10^5n^2 + 10^8n$ 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

\[
f(n) \leq cg(n) \quad \text{for} \quad n \geq n_0
\]

Example: \( 2n + 10 \) is \( O(n) \)

For what \( c \) and \( n \) is \( cg(n) \geq f(n) \)?

- \( 2n + 10 \leq cn \)
- \( (c - 2) n \geq 10 \)
- \( n \geq 10/(c - 2) \)
- 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

  e.g. Is there a \(c\) and \(n_0\) such that \(c \cdot n \geq 7n-2\)?

  What is the big-Oh?
  
  7n-2 is \(O(n)\)

  [TEAMS] What are the values for \(c\) and \(n_0\)?
  
  Need \(c > 0\) and \(n_0 \geq 1\) such that \(7n-2 \leq c \cdot n\) for \(n \geq n_0\)
  
  This is true for \(c = 7\) and \(n_0 = 1\)
More Big-Oh Examples

- $2n^2 + 16$
  
  $2n^2 + 16$ is $O(n^2)$
  
  need $c > 0$ and $n_0 \geq 1$ such that $2n^2 + 16 \leq c \cdot n^2$ for $n \geq n_0$
  
  this is true for $c = 3$ and $n_0 = 4$

- $3n^3 + 2n^2 + 9$
  
  $3n^3 + 2n^2 + 9$ is $O(n^3)$
  
  need $c > 0$ and $n_0 \geq 1$ such that $3n^3 + 2n^2 + 9 \leq c \cdot n^3$ for $n \geq n_0$
  
  this is true for $c = 4$ and $n_0 = 3$
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.

<table>
<thead>
<tr>
<th>( g(n) ) grows more</th>
<th>( f(n) ) is ( O(g(n)) )</th>
<th>( g(n) ) is ( O(f(n)) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Same growth</th>
<th>( f(n) ) grows more</th>
<th>( g(n) ) grows more</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Big-Oh Rules

If \( f(n) \) is 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 determine that algorithm `arrayMax` executes at most $7n - 1$ primitive operations.
  - 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.
Computing Prefix Averages

- We further illustrate asymptotic analysis with two algorithms for prefix averages.
- The $i$-th prefix average of an array $X$ is average of the first $(i+1)$ elements of $X$:
  \[ A[i] = \frac{(X[0] + X[1] + \ldots + X[i])}{(i+1)} \]
- Computing the array $A$ of prefix averages of another array $X$ has applications to financial analysis.
Prefix Averages (Quadratic)

The following algorithm computes prefix averages in quadratic time by applying the definition.

Algorithm `prefixAverages1(X, n)`

Input array `X` of `n` integers
Output array `A` of prefix averages of `X`

```plaintext
A ← new array of `n` integers
for `i` ← 0 to `n` - 1 do
    `s` ← `X[0]`
    for `j` ← 1 to `i` do
        `s` ← `s + X[j]`
    `A[i]` ← `s / (i + 1)`
return `A`
```

#operations
`n`
`n`
`n`
`1 + 2 + … + (n - 1)`
`1 + 2 + … + (n - 1)`
`n`
`1`
Arithmetic Progression

- The running time of `prefixAverages1` 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 `prefixAverages1` runs in $O(n^2)$ time
Prefix Averages

Asymptotic analysis tells us the complexity of the original algorithm.

[TEAMS] Given such a tool, how can we make the algorithm more efficient?
Prefix Averages (Linear)

The following algorithm computes prefix averages in linear time by keeping a running sum:

**Algorithm** `prefixAverages2(X, n)`

**Input** array `X` of `n` integers

**Output** array `A` of prefix averages of `X`

1. `A ← new array of n integers`
2. `s ← 0`
3. for `i ← 0` to `n − 1` do
   1. `s ← s + X[i]`
   2. `A[i] ← s / (i + 1)`
4. return `A`

Algorithm `prefixAverages2` runs in $O(n)$ time.
Math you need to Review

- Summations
- Logarithms and Exponents
- Proof techniques
- Basic probability

**properties of logarithms:**
\[
\log_b(xy) = \log_b x + \log_b y
\]
\[
\log_b (x/y) = \log_b x - \log_b y
\]
\[
\log_b xa = a \log_b x
\]
\[
\log_b a = \log_x a / \log_x b
\]

**properties of exponentials:**
\[
a^{(b+c)} = a^b a^c
\]
\[
a^{bc} = (a^b)^c
\]
\[
a^b / a^c = a^{(b-c)}
\]
\[
b = a^{\log_a b}
\]
\[
b^c = a^{c \log_a b}
\]
Relatives of Big-Oh

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

- **big-Theta**
  - $f(n)$ is $\Theta(g(n))$ if there are constants $c' > 0$ and $c'' > 0$ and an integer constant $n_0 \geq 1$ such that $c'g(n) \leq f(n) \leq c''g(n)$ for $n \geq n_0$

- **little-o**
  - $f(n)$ is $o(g(n))$ if, for any constant $c > 0$, there is an integer constant $n_0 \geq 0$ such that $f(n) \leq c \cdot g(n)$ for $n \geq n_0$

- **little-omega**
  - $f(n)$ is $\omega(g(n))$ if, for any constant $c > 0$, there is an integer constant $n_0 \geq 0$ such that $f(n) \geq c \cdot g(n)$ for $n \geq n_0$
Intuition for Asymptotic Notation

**Big-Oh**
- $f(n)$ is $O(g(n))$ if $f(n)$ is asymptotically \textbf{less than or equal} to $g(n)$

**big-Omega**
- $f(n)$ is $\Omega(g(n))$ if $f(n)$ is asymptotically \textbf{greater than or equal} to $g(n)$

**big-Theta**
- $f(n)$ is $\Theta(g(n))$ if $f(n)$ is asymptotically \textbf{equal} to $g(n)$

**little-o**
- $f(n)$ is $o(g(n))$ if $f(n)$ is asymptotically \textbf{strictly less} than $g(n)$

**little-omega**
- $f(n)$ is $\omega(g(n))$ if is asymptotically \textbf{strictly greater} than $g(n)$
Example Uses of the Relatives of Big-Oh

■ \(5n^2\) is \(\Omega(n^2)\)

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

[TEAMS] What are values for \(c\) and \(n_0\)?

Let \(c = 5\) and \(n_0 = 1\)

■ \(5n^2\) is \(\Omega(n)\)

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

Let \(c = 1\) and \(n_0 = 1\)

■ \(5n^2\) is \(\omega(n)\)

\[ f(n) = \omega(g(n)) \] if, for any constant \(c > 0\), there is an integer constant \(n_0 \geq 0\) such that \(f(n) \geq c \cdot g(n)\) for \(n \geq n_0\)

Need \(5n_0^2 \geq c \cdot n_0 \rightarrow\) given \(c\), the \(n_0\) that satisfies this is \(n_0 \geq c/5 \geq 0\)
Euclid’s Algorithm

An algorithm for computing the greatest common divisor (GCD) of two numbers $M \geq N$:

Algorithm GCD($M, N$)

while ($N \neq 0$)

    rem $\leftarrow M \mod N$
    $M \leftarrow N$
    $N \leftarrow \text{rem}$

endwhile

return $M$

[TEAMS] What is the Big-Oh?
Euclid’s Algorithm

An algorithm for computing the greatest common divisor (GCD) of two numbers $M \geq N$:

Algorithm GCD(M, N)

while $(N \neq 0)$

rem ← $M \mod N$

$M ← N$

$N ← \text{rem}$

endwhile

return $M$

The algorithm GCD runs is $O(\log n)$
Euclid’s Algorithm

Why?

- If $M \geq N$, then $(M \mod N) < M/2$
- Thus, each iteration at least halves the value of $M$
Exponentiation

A recursive algorithm to compute $X$ to the power $N$:

Algorithm pow(X, N)
    if (N == 0) return 1
    if (N == 1) return X
    if (N is even)
        return pow(X*X, N/2)
    else
        return pow(X*X, N/2) * X

[TEAMS] What is the Big-Oh?
Exponentiation

A recursive algorithm to compute $X$ to the power $N$:

Algorithm `pow(X, N)`

- if $(N == 0)$ return 1
- if $(N == 1)$ return $X$
- if $(N$ is even)
  - return `pow(X*X, N/2)`
- else
  - return `pow(X*X, N/2) * X`

The algorithm `pow` is $O(\log n)$