CS 5114: Theory of Algorithms

Clifford A. Shaffer

Department of Computer Science
Virginia Tech
Blacksburg, Virginia

Spring 2010

Title page

CS5114: Theory of Algorithms

- Emphasis: Creation of Algorithms
- Less important:
  - Analysis of algorithms
  - Problem statement
  - Programming
- Central Paradigm: Mathematical Induction
  - Find a way to solve a problem by solving one or more smaller problems

Creation of algorithms comes through exploration, discovery, techniques, intuition: largely by **lots** of examples and **lots** of practice (HW exercises).
We will use Analysis of Algorithms as a tool.
Problem statement (in the software eng. sense) is not important because our problems are easily described, if not easily solved.
Smaller problems may or may not be the same as the original problem.
Divide and conquer is a way of solving a problem by solving one more more smaller problems.
Claim on induction: The processes of constructing proofs and constructing algorithms are similar.

Review of Mathematical Induction

- The paradigm of **Mathematical Induction** can be used to solve an enormous range of problems.
- **Purpose**: To prove a parameterized theorem of the form:
  Theorem: $\forall n \geq c, \mathbf{P}(n)$.
  - Use only positive integers $\geq c$ for n.
- Sample $\mathbf{P}(n)$:
  $n + 1 \leq n^2$

$\mathbf{P}(n)$ is a statement containing n as a variable.

This sample $\mathbf{P}(n)$ is true for $n \geq 2$, but false for $n = 1$.

Principle of Mathematical Induction

- IF the following two statements are true:
  1. $\mathbf{P}(c)$ is true.
  2. For $n > c, \mathbf{P}(n-1)$ is true $\rightarrow \mathbf{P}(n)$ is true.

  ... **THEN** we may conclude: $\forall n \geq c$, $\mathbf{P}(n)$.
- The assumption "$\mathbf{P}(n-1)$ is true" is the **induction hypothesis**.
- Typical induction proof form:
  1. Base case
  2. State induction Hypothesis
  3. Prove the implication (induction step)
- What does this remind you of?

Important: The goal is to prove the **implication**, not the theorem! That is, prove that $\mathbf{P}(n-1) \rightarrow \mathbf{P}(n)$. **NOT** to prove $P(n)$. This is much easier, because we can assume that $\mathbf{P}(n)$ is true.
Consider the truth table for implication to see this. Since $A \rightarrow B$ is (vacuously) true when A is false, we can just assume that A is true since the implication is true anyway A is false. That is, we only need to worry that the implication could be false if A is true.

The power of induction is that the induction hypothesis "comes for free." We often try to make the most of the extra information provided by the induction hypothesis.
This is like recursion! There you have a base case and a recursive call that must make progress toward the base case.

Induction Example 1

Theorem: Let

$$S(n) = \sum_{i=1}^{n} i = 1 + 2 + \cdots + n.$$

Then, $\forall n \geq 1, S(n) = \frac{n(n+1)}{2}$.

Induction Example 1
Theorem: Let
$S(n) = \sum_{i=1}^{n} i = 1 + 2 + \cdots + n$.
Then, $\forall n \geq 1, S(n) = \frac{n(n+1)}{2}$.

Base Case: $\mathbf{P}(n)$ is true since $S(1) = 1 = 1(1+1)/2$.
Induction Hypothesis: $S(i) = \frac{i(i+1)}{2}$ for $i < n$.
Induction Step:

$$\begin{aligned} S(n) &= S(n-1) + n = (n-1)n/2 + n \\ &= \frac{n(n+1)}{2} \end{aligned}$$

Therefore, $\mathbf{P}(n-1) \rightarrow \mathbf{P}(n)$.
By the principle of Mathematical Induction,
$\forall n \geq 1, S(n) = \frac{n(n+1)}{2}$.
MI is often an ideal tool for **verification** of a hypothesis.
Unfortunately it does not help to construct a hypothesis.

Induction Example 2

Theorem: $\forall n \geq 1, \forall$ real x such that $1 + x > 0$,
$(1+x)^n \geq 1 + nx$.

Induction Example 2
Theorem: $\forall n \geq 1, \forall$ real x such that $1 + x > 0$,
$(1+x)^n \geq 1 + nx$.

What do we do induction on? Can't be a real number, so must be n.
$\mathbf{P}(n) : (1+x)^n \geq 1 + nx$.

Base Case: $(1+x)^1 = 1 + x \geq 1 + 1x$
Induction Hypothesis: Assume $(1+x)^{n-1} \geq 1 + (n-1)x$
Induction Step:

$$\begin{aligned} (1+x)^n &= (1+x)(1+x)^{n-1} \\ &\geq (1+x)(1+(n-1)x) \\ &= 1 + nx - x + x + nx^2 - x^2 \\ &= 1 + nx + (n-1)x^2 \\ &\geq 1 + nx. \end{aligned}$$

Induction Example 3

Theorem: 2¢ and 5¢ stamps can be used to form any denomination (for denominations ≥ 4).

Induction Example 3
Theorem: 2¢ and 5¢ stamps can be used to form any denomination (for denominations ≥ 4).

Base case: $4 = 2 + 2$.

Induction Hypothesis: Assume $\mathbf{P}(k)$ for $4 \leq k < n$.

Induction Step:
Case 1: $n - 1$ is made up of all 2¢ stamps. Then, replace 2 of these with a 5¢ stamp.

Case 2: $n - 1$ includes a 5¢ stamp. Then, replace this with 3 2¢ stamps.

Colorings

4-color problem: For any set of polygons, 4 colors are sufficient to guarentee that no two adjacent polygons share the same color.

Restrict the problem to regions formed by placing (infinite) lines in the plane. How many colors do we need?
Candidates:

- 4: Certainly
- 3: ?
- 2: ?
- 1: No!

Let's try it for 2...

Colorings
4-color problem: For any set of polygons, 4 colors are sufficient to guarentee that no two adjacent polygons share the same color.
Restrict the problem to regions formed by placing (infinite) lines in the plane. How many colors do we need?
Candidates:
4: Certainly
3: ?
2: ?
1: No!
Let's try it for 2...

Induction is useful for much more than checking equations!

If we accept the statement about the general 4-color problem, then of course 4 colors is enough for our restricted version.

If 2 is enough, then of course we can do it with 3 or more.

Two-coloring Problem

Given: Regions formed by a collection of (infinite) lines in the plane.
Rule: Two regions that share an edge cannot be the same color.

Theorem: It is possible to two-color the regions formed by n lines.

Two-coloring Problem

Given: Regions formed by a collection of (infinite) lines in the plane.
Rule: Two regions that share an edge cannot be the same color.

Theorem: It is possible to two-color the regions formed by n lines.

Picking what to do induction on can be a problem. Lines? Regions? How can we "add a region?" We can't, so try induction on lines.
Base Case: $n = 1$. Any line divides the plane into two regions.
Induction Hypothesis: It is possible to two-color the regions formed by $n - 1$ lines.
Induction Step: Introduce the n'th line.
This line cuts some colored regions in two.
Reverse the region colors on one side of the n'th line.
A valid two-coloring results.

- Any boundary surviving the addition still has opposite colors.
- Any new boundary also has opposite colors after the switch.

Strong Induction

IF the following two statements are true:

1. $\mathbf{P}(c)$
2. $\mathbf{P}(i), i = 1, 2, \cdots, n-1 \rightarrow \mathbf{P}(n)$,

... **THEN** we may conclude: $\forall n \geq c$, $\mathbf{P}(n)$.

Advantage: We can use statements other than $\mathbf{P}(n-1)$ in proving $\mathbf{P}(n)$.

Strong Induction

IF the following two statements are true:
1. $\mathbf{P}(c)$
2. $\mathbf{P}(i), i = 1, 2, \cdots, n-1 \rightarrow \mathbf{P}(n)$,

... **THEN** we may conclude: $\forall n \geq c$, $\mathbf{P}(n)$.

Advantage: We can use statements other than $\mathbf{P}(n-1)$ in proving $\mathbf{P}(n)$.

The previous examples were all very straightforward – simply add in the n'th item and justify that the IH is maintained.
Now we will see examples where we must do more sophisticated (creative!) maneuvers such as

- go backwards from n.
- prove a stronger IH.

to make the most of the IH.

Graph Problem

An **Independent Set** of vertices is one for which no two vertices are adjacent.

Theorem: Let $G = (V, E)$ be a **directed** graph. Then, G contains some independent set $S(G)$ such that every vertex can be reached from a vertex in $S(G)$ by a path of length at most 2.

Example: a graph with 3 vertices in a cycle. Pick any one vertex as $S(G)$.

Graph Problem

An **Independent Set** of vertices is one for which no two vertices are adjacent.

Theorem: Let $G = (V, E)$ be a **directed** graph. Then, G contains some independent set $S(G)$ such that every vertex can be reached from a vertex in $S(G)$ by a path of length at most 2.

Example: a graph with 3 vertices in a cycle. Pick any one vertex as $S(G)$.

It should be obvious that the theorem is true for an undirected graph.
Naive approach: Assume the theorem is true for any graph of $n - 1$ vertices. Now add the nth vertex and its edges. But this won't work for the graph $1 \leftarrow 2$. Initially, vertex 1 is the independent set. We can't add 2 to the graph. Nor can we reach it from 1.
Going forward is good for proving existance.
Going backward (from an arbitrary instance into the IH) is usually necessary to prove that a property holds in all instances. This is because going forward requires proving that you reach all of the possible instances.

Graph Problem (cont)

Theorem: Let $G = (V, E)$ be a **directed** graph. Then, G contains some independent set $S(G)$ such that every vertex can be reached from a vertex in $S(G)$ by a path of length at most 2.
Base Case: Easy if $n \leq 3$ because there can be no path of length > 2.
Induction Hypothesis: The theorem is true if $|V| < n$.
Induction Step ($n > 3$):
Pick any $v \in V$.
Define: $N(v) = \{v\} \cup \{w \in V | (v, w) \in E\}$.
$H = G - N(v)$.
Since the number of vertices in H is less than n, there is an independent set $S(H)$ that satisfies the theorem for H.

Graph Problem (cont)

Theorem: Let $G = (V, E)$ be a **directed** graph. Then, G contains some independent set $S(G)$ such that every vertex can be reached from a vertex in $S(G)$ by a path of length at most 2.
Base Case: Easy if $n \leq 3$ because there can be no path of length > 2.
Induction Hypothesis: The theorem is true if $|V| < n$.
Induction Step ($n > 3$):
Pick any $v \in V$.
Define: $N(v) = \{v\} \cup \{w \in V | (v, w) \in E\}$.
$H = G - N(v)$.
Since the number of vertices in H is less than n, there is an independent set $S(H)$ that satisfies the theorem for H.

$N(v)$ is all vertices reachable (directly) from v. That is, the Neighbors of v.
H is the graph induced by $V - N(v)$.

OK, so why remove both v and $N(v)$ from the graph? If we only remove v, we have the same problem as before. If G is $1 \rightarrow 2 \rightarrow 3$, and we remove 1, then the independent set for H must be vertex 2. We can't just add back 1. But if we remove both 1 and 2, then we'll be able to do something...

Graph Proof (cont)

There are two cases:

1. $S(H) \cup \{v\}$ is independent.
   Then $S(G) = S(H) \cup \{v\}$.
2. $S(H) \cup \{v\}$ is not independent.
   Let $w \in S(H)$ such that $(w, v) \in E$.
   Every vertex in $N(v)$ can be reached by w with path of length ≤ 2.
   So, set $S(G) = S(H)$.

By Strong Induction, the theorem holds for all G.

Graph Proof (cont)

There are two cases:
1. $S(H) \cup \{v\}$ is independent.
   Then $S(G) = S(H) \cup \{v\}$.
2. $S(H) \cup \{v\}$ is not independent.
   Let $w \in S(H)$ such that $(w, v) \in E$.
   Every vertex in $N(v)$ can be reached by w with path of length ≤ 2.
   So, set $S(G) = S(H)$.

By Strong Induction, the theorem holds for all G.

"$S(H) \cup \{v\}$ is not independent" means that there is an edge from something in $S(H)$ to v.
IMPORTANT: There cannot be an edge from v to $S(H)$ because whatever we can reach from v is in $N(v)$ and would have been removed in H.
We need strong induction for this proof because we don't know how many vertices are in $N(v)$.

Fibonacci Numbers

Define Fibonacci numbers inductively as:

$$\begin{aligned} F(1) &= F(2) = 1 \\ F(n) &= F(n-1) + F(n-2), n > 2. \end{aligned}$$

Theorem: $\forall n \geq 1, F(n)^2 + F(n+1)^2 = F(2n+1)$.

Induction Hypothesis:
$F(n-1)^2 + F(n)^2 = F(2n-1)$.

Fibonacci Numbers

Define Fibonacci numbers inductively as:

$$\begin{aligned} F(1) &= F(2) = 1 \\ F(n) &= F(n-1) + F(n-2), n > 2. \end{aligned}$$

Theorem: $\forall n \geq 1, F(n)^2 + F(n+1)^2 = F(2n+1)$.

Induction Hypothesis:
$F(n-1)^2 + F(n)^2 = F(2n-1)$.

Expand both sides of the theorem, then cancel like terms:
$F(2n+1) = F(2n) + F(2n-1)$ and,

$$\begin{aligned} F(n)^2 + F(n+1)^2 &= F(n)^2 + (F(n) + F(n-1))^2 \\ &= F(n)^2 + F(n)^2 + 2F(n)F(n-1) + F(n-1)^2 \\ &= F(n)^2 + F(n-1)^2 + F(n)^2 + 2F(n)F(n-1) \\ &= F(2n-1) + F(n)^2 + 2F(n)F(n-1). \end{aligned}$$

Want: $F(n)^2 + F(n+1)^2 = F(2n+1) = F(2n) + F(2n-1)$
Steps above gave:
$F(2n) + F(2n-1) = F(2n-1) + F(n)^2 + 2F(n)F(n-1)$
So we need to show that: $F(n)^2 + 2F(n)F(n-1) = F(2n)$
To prove the original theorem, we must prove this. Since we must do it anyway, we should take advantage of this in our IH!

Fibonacci Numbers (cont)

With a stronger theorem comes a stronger IH!

Theorem:
$F(n)^2 + F(n+1)^2 = F(2n+1)$ and
$F(n)^2 + 2F(n)F(n-1) = F(2n)$.

Induction Hypothesis:
$F(n-1)^2 + F(n)^2 = F(2n-1)$ and
$F(n-1)^2 + 2F(n-1)F(n-2) = F(2n-2)$.

Fibonacci Numbers (cont)

With a stronger theorem comes a stronger IH!

Theorem:
$F(n)^2 + F(n+1)^2 = F(2n+1)$ and
$F(n)^2 + 2F(n)F(n-1) = F(2n)$.

Induction Hypothesis:
$F(n-1)^2 + F(n)^2 = F(2n-1)$ and
$F(n-1)^2 + 2F(n-1)F(n-2) = F(2n-2)$.

$F(n)^2 + 2F(n)F(n-1)$

$$\begin{aligned} &= F(n)^2 + 2(F(n-1) + F(n-2))F(n-1) \\ &= F(n)^2 + F(n-1)^2 + 2F(n-1)F(n-2) + F(n-1)^2 \\ &= F(2n-1) + F(2n-2) \\ &= F(2n). \end{aligned}$$

$$\begin{aligned} F(n)^2 + F(n+1)^2 &= F(n)^2 + [F(n) + F(n-1)]^2 \\ &= F(n)^2 + F(n)^2 + 2F(n)F(n-1) + F(n-1)^2 \\ &= F(n)^2 + F(2n) + F(n-1)^2 \\ &= F(2n-1) + F(2n) \\ &= F(2n+1). \end{aligned}$$

... which proves the theorem. The original result could not have been proved without the stronger induction hypothesis.

Another Example

Theorem: All horses are the same color.

Proof: $\mathbf{P}(n)$: If S is a set of n horses, then all horses in S have the same color.
Base case: $n = 1$ is easy.
Induction Hypothesis: Assume $\mathbf{P}(i), i < n$.
Induction Step:

- Let S be a set of horses, $|S| = n$.
- Let S' be $S - \{h\}$ for some horse h.
- By IH, all horses in S' have the same color.
- Let h' be some horse in S'.
- IH implies $\{h, h'\}$ have all the same color.

Therefore, $\mathbf{P}(n)$ holds.

Another Example

Theorem: All horses are the same color.

Proof: $\mathbf{P}(n)$: If S is a set of n horses, then all horses in S have the same color.
Base case: $n = 1$ is easy.
Induction Hypothesis: Assume $\mathbf{P}(i), i < n$.
Induction Step:
- Let S be a set of horses, $|S| = n$.
- Let S' be $S - \{h\}$ for some horse h.
- By IH, all horses in S' have the same color.
- Let h' be some horse in S'.
- IH implies $\{h, h'\}$ have all the same color.

Therefore, $\mathbf{P}(n)$ holds.

The problem is that the base case does not give enough strength to give the **particular** instance of $n = 2$ used in the last step.

Algorithm Analysis

- We want to "measure" algorithms.
- What do we measure?
- What factors affect measurement?
- Objective: Measures that are independent of all factors except input.

Algorithm Analysis
- We want to "measure" algorithms.
- What do we measure?
- What factors affect measurement?
- Objective: Measures that are independent of all factors except input.

What do we measure?
Time and space to run; ease of implementation (this changes with language and tools); code size

What affects measurement?
Computer speed and architecture; Programming language and compiler; System load; Programmer skill; Specifics of input (size, arrangement)

If you compare two programs running on the same computer under the same conditions, all the other factors (should) cancel out.
Want to measure the relative efficiency of two algorithms without needing to implement them on a real computer.

Time Complexity

- Time and space are the most important computer resources.
- Function of input: **T**(input)
- Growth of time with size of input:
  - Establish an (integer) **<u>size</u>** n for inputs
  - n numbers in a list
  - n edges in a graph
- Consider time for all inputs of size n:
  - Time varies widely with specific input
  - Best case
  - Average case
  - Worst case
- Time complexity $\mathbf{T}(n)$ counts **<u>steps</u>** in an algorithm.

Time Complexity
- Time and space are the most important computer resources.
- Function of input: T(input)
- Growth of time with size of input:
  - Establish an (integer) size n for inputs
  - n numbers in a list
  - n edges in a graph
- Consider time for all inputs of size n:
  - Time varies widely with specific input
  - Best case
  - Average case
  - Worst case
- Time complexity T(n) counts steps in an algorithm.

Sometimes analyze in terms of more than one variable.
Best case usually not of interest.
Average case is usually what we want, but can be hard to measure.
Worst case appropriate for "real-time" applications, often best we can do in terms of measurement.
Examples of "steps:" comparisons, assignments, arithmetic/logical operations. What we choose for "step" depends on the algorithm. Step cost must be "constant" – not dependent on n.

Asymptotic Analysis

- It is undesirable/impossible to count the exact number of steps in most algorithms.
  - Instead, concentrate on main characteristics.
- Solution: Asymptotic analysis
  - Ignore small cases:
    - Consider behavior approaching infinity
  - Ignore constant factors, low order terms:
    - $2n^2$ looks the same as $5n^2 + n$ to us.

Asymptotic Analysis
- It is undesirable/impossible to count the exact number of steps in most algorithms.
  - Instead, concentrate on main characteristics.
- Solution: Asymptotic analysis
  - Ignore small cases:
    - Consider behavior approaching infinity
  - Ignore constant factors, low order terms:
    - $2n^2$ looks the same as $5n^2 + n$ to us.

Undesirable to count number of machine instructions or steps because issues like processor speed muddy the waters.

O Notation

O notation is a measure for "upper bound" of a growth rate.
- pronounced "Big-oh"

Definition: For $\mathbf{T}(n)$ a non-negatively valued function, $\mathbf{T}(n)$ is in the set $\mathrm{O}(f(n))$ if there exist two positive constants c and n_0 such that $\mathbf{T}(n) \leq cf(n)$ for all $n > n_0$.

Examples:
- $5n + 8 \in \mathrm{O}(n)$
- $2n^2 + n \log n \in \mathrm{O}(n^2) \in \mathrm{O}(n^3 + 5n^2)$
- $2n^2 + n \log n \in \mathrm{O}(n^2) \in \mathrm{O}(n^3 + n^2)$

O Notation
O notation is a measure for "upper bound" of a growth rate.
- pronounced "Big-oh"
Definition: For T(n) a non-negatively valued function, T(n) is in the set O(f(n)) if there exist two positive constants c and n_0 such that T(n) ≤ cf(n) for all n > n_0.
Examples:
- 5n + 8 ∈ O(n)
- $2n^2 + n \log n \in O(n^2) \in O(n^3 + 5n^2)$
- $2n^2 + n \log n \in O(n^2) \in O(n^3 + n^2)$

Remember: The time equation is for some particular set of inputs – best, worst, or average case.

O Notation (cont)

We seek the "simplest" and "strongest" f.

Big-O is somewhat like "$\leq$":
$n^2 \in \mathrm{O}(n^3)$ and $n^2 \log n \in O(n^3)$, but

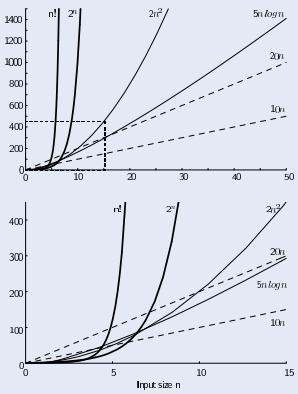
- $n^2 \neq n^2 \log n$
- $n^2 \in \mathrm{O}(n^2)$ while $n^2 \log n \notin \mathrm{O}(n^2)$

O Notation (cont)
We seek the "simplest" and "strongest" f.
Big-O is somewhat like "≤":
n² ∈ O(n³) and n² log n ∈ O(n³), but
n² ≠ n² log n
n² ∈ O(n²) while n² log n ∉ O(n²)

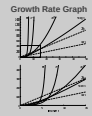
A common misunderstanding:

- "The best case for my algorithm is $n = 1$ because that is the fastest." WRONG!
- Big-oh refers to a growth rate as n grows to ∞.
- Best case is defined for the input of size n that is cheapest among all inputs of size n.

Growth Rate Graph



Growth Rate Graph



2^n is an exponential algorithm. $10n$ and $20n$ differ only by a constant.

Speedups

What happens when we buy a computer 10 times faster?

$\mathbf{T}(n)$	n	n'	Change	n'/n
$10n$	$1,000$	$10,000$	$n' = 10n$	10
$20n$	500	$5,000$	$n' = 10n$	10
$5n \log n$	250	$1,842$	$\sqrt{10}n < n' < 10n$	7.37
$2n^2$	70	223	$n' = \sqrt{10}n$	3.16
2^n	13	16	$n' = n + 3$	$--$

n: Size of input that can be processed in one hour (10,000 steps).

n': Size of input that can be processed in one hour on the new machine (100,000 steps).

Speedups
What happens when we buy a computer 10 times faster?
T(n) | n | n′ | Change | n′/n
10n | 1,000 | 10,000 | n′ = 10n | 10
20n | 500 | 5,000 | n′ = 10n | 10
5nlogn | 250 | 1,842 | √10n < n′ < 10n | 7.37
2n² | 70 | 223 | n′ = √10n | 3.16
2ⁿ | 13 | 16 | n′ = n + 3 | −−
n: Size of input that can be processed in one hour (10,000 steps).
n′: Size of input that can be processed in one hour on the new machine (100,000 steps).

How much speedup? 10 times. More important: How much increase in problem size for same time? Depends on growth rate.
For n^2, if $n = 1000$, then n' would be 1003.
Compare $\mathbf{T}(n) = n^2$ to $\mathbf{T}(n) = n \log n$. For $n > 58$, it is faster to have the $\Theta(n \log n)$ algorithm than to have a computer that is 10 times faster.

Some Rules for Use

Definition: f is <u>**monotonically growing**</u> if $n_1 \geq n_2$ implies $f(n_1) \geq f(n_2)$.
We typically assume our time complexity function is monotonically growing.

Theorem 3.1: Suppose f is monotonically growing.
$\forall c > 0$ and $\forall a > 1, (f(n))^c \in O(a^{f(n)})$
In other words, an <u>**exponential**</u> function grows faster than a <u>**polynomial**</u> function.

Lemma 3.2: If $f(n) \in O(s(n))$ and $g(n) \in O(r(n))$ then

- $f(n) + g(n) \in O(s(n) + r(n)) \equiv O(\max(s(n), r(n)))$
- $f(n)g(n) \in O(s(n)r(n))$.
- If $s(n) \in O(h(n))$ then $f(n) \in O(h(n))$
- For any constant k, $f(n) \in O(ks(n))$

Some Rules for Use
Definition: f is monotonically growing if n₁ ≥ n₂ implies f(n₁) ≥ f(n₂).
We typically assume our time complexity function is monotonically growing.
Theorem 3.1: Suppose f is monotonically growing.
∀c > 0 and ∀a > 1, (f(n))ᶜ ∈ O(aᶠ⁽ⁿ⁾)
In other words, an exponential function grows faster than a polynomial function.
Lemma 3.2: If f(n) ∈ O(s(n)) and g(n) ∈ O(r(n)) then
f(n) + g(n) ∈ O(s(n) + r(n)) ≡ O(max(s(n), r(n)))
f(n)g(n) ∈ O(s(n)r(n)).
If s(n) ∈ O(h(n)) then f(n) ∈ O(h(n))
For any constant k, f(n) ∈ O(ks(n))

Assume monitonic growth because larger problems should take longer to solve. However, many real problems have "cyclically growing" behavior.
Is $O(2^{f(n)}) \in O(3^{f(n)})$? Yes, but not vice versa.
$3^n = 1.5^n \times 2^n$ so no constant could ever make 2^n bigger than 3^n for all n.functional composition

Other Asymptotic Notation

$\Omega(f(n))$ – lower bound ($\geq$)
Definition: For $\mathbf{T}(n)$ a non-negatively valued function, $\mathbf{T}(n)$ is in the set $\Omega(g(n))$ if there exist two positive constants c and n_0 such that $\mathbf{T}(n) \geq cg(n)$ for all $n > n_0$.
Ex: $n^2 \log n \in \Omega(n^2)$.

$\Theta(f(n))$ – Exact bound ($=$)
Definition: $g(n) = \Theta(f(n))$ if $g(n) \in O(f(n))$ and $g(n) \in \Omega(f(n))$.
Important!: It is Θ if it is both in big-Oh and in Ω.
Ex: $5n^3 + 4n^2 + 9n + 7 = \Theta(n^3)$

Other Asymptotic Notation

$\Omega(f(n))$ – lower bound ($\geq$)
Definition: For $\mathbf{T}(n)$ a non-negatively valued function, $\mathbf{T}(n)$ is in the set $\Omega(g(n))$ if there exist two positive constants c and n_0 such that $\mathbf{T}(n) \geq cg(n)$ for all $n > n_0$.
Ex: $n^2 \log n \in \Omega(n^2)$.

$\Theta(f(n))$ – Exact bound ($=$)
Definition: $g(n) = \Theta(f(n))$ if $g(n) \in O(f(n))$ and $g(n) \in \Omega(f(n))$.
Important: It is Θ if it is both in big-Oh and in Ω.
Ex: $5n^3 + 4n^2 + 9n + 7 = \Theta(n^3)$

Ω is most userful to discuss cost of problems, not algorithms. Once you have an equation, the bounds have met. So this is more interesting when discussing your level of uncertainty about the difference between the upper and lower bound.

You have Θ when you have the upper and the lower bounds meeting. So Θ means that you know a lot more than just Big-oh, and so is perferred when possible.

A common misunderstanding:

- Confusing worst case with upper bound.
- Upper bound refers to a growth rate.
- Worst case refers to the worst input from among the choices for possible inputs of a given size.

Other Asymptotic Notation (cont)

$o(f(n))$ – little o ($<$)
Definition: $g(n) \in o(f(n))$ if $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$
Ex: $n^2 \in o(n^3)$

$\omega(f(n))$ – little omega ($>$)
Definition: $g(n) \in w(f(n))$ if $f(n) \in o(g(n))$.
Ex: $n^5 \in w(n^2)$

$\infty(f(n))$
Definition: $T(n) = \infty(f(n))$ if $T(n) = O(f(n))$ but the constant in the O is so large that the algorithm is impractical.

Other Asymptotic Notation (cont)

$o(f(n))$ – little o ($<$)
Definition: $g(n) \in o(f(n))$ if $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$
Ex: $n^2 \in o(n^3)$

$\omega(f(n))$ – little omega ($>$)
Definition: $g(n) \in w(f(n))$ if $f(n) \in o(g(n))$.
Ex: $n^5 \in w(n^2)$

$\infty(f(n))$
Definition: $T(n) = \infty(f(n))$ if $T(n) = O(f(n))$ but the constant in the O is so large that the algorithm is impractical.

We won't use these too much.

Aim of Algorithm Analysis

Typically want to find "simple" $f(n)$ such that $T(n) = \Theta(f(n))$.

- Sometimes we settle for $O(f(n))$.

Usually we measure T as "worst case" time complexity.
Sometimes we measure "average case" time complexity.
Approach: Estimate number of "steps"

- Appropriate step depends on the problem.
- Ex: measure key comparisons for sorting

Summation: Since we typically count steps in different parts of an algorithm and sum the counts, techniques for computing sums are important (loops).

Recurrence Relations: Used for counting steps in recursion.

Aim of Algorithm Analysis

Typically want to find "simple" $f(n)$ such that $T(n) = \Theta(f(n))$.
- Sometimes we settle for $O(f(n))$.

Usually we measure T as "worst case" time complexity.
Sometimes we measure "average case" time complexity.
Approach: Estimate number of "steps"
- Appropriate step depends on the problem.
- Ex: measure key comparisons for sorting

Summation: Since we typically count steps in different parts of an algorithm and sum the counts, techniques for computing sums are important (loops).

Recurrence Relations: Used for counting steps in recursion.

We prefer Θ over Big-oh because Θ means that we understand our bounds and they met. But if we just can't find that the bottom meets the top, then we are stuck with just Big-oh. Lower bounds can be hard. For **problems** we are often interested in Ω

– but this is often hard for non-trivial situations!

Often prefer average case (except for real-time programming), but worst case is simpler to compute than average case since we need not be concerned with distribution of input.

For the sorting example, key comparisons must be constant-time to be used as a cost measure.

Summation: Guess and Test

Technique 1: Guess the solution and use induction to test.

Technique 1a: Guess the form of the solution, and use simultaneous equations to generate constants. Finally, use induction to test.

Summation: Guess and Test

Technique 1: Guess the solution and use induction to test.

Technique 1a: Guess the form of the solution, and use simultaneous equations to generate constants. Finally, use induction to test.

no notes

Summation Example

$$S(n) = \sum_{i=0}^{n} i^2.$$

Guess that $S(n)$ is a polynomial $\leq n^3$.
Equivalently, guess that it has the form
$S(n) = an^3 + bn^2 + cn + d$.

For $n = 0$ we have $S(n) = 0$ so $d = 0$.
For $n = 1$ we have $a + b + c + 0 = 1$.
For $n = 2$ we have $8a + 4b + 2c = 5$.
For $n = 3$ we have $27a + 9b + 3c = 14$.

Solving these equations yields $a = \frac{1}{3}, b = \frac{1}{2}, c = \frac{1}{6}$

Now, prove the solution with induction.

Summation Example
$S(n) = \sum_{i=0}^{n} i^2$.
Guess that $S(n)$ is a polynomial $\leq n^3$.
Equivalently, guess that it has the form
$S(n) = an^3 + bn^2 + cn + d$.
For $n = 0$ we have $S(n) = 0$ so $d = 0$.
For $n = 1$ we have $a + b + c + 0 = 1$.
For $n = 2$ we have $8a + 4b + 2c = 5$.
For $n = 3$ we have $27a + 9b + 3c = 14$.
Solving these equations yields $a = \frac{1}{3}, b = \frac{1}{2}, c = \frac{1}{6}$
Now, prove the solution with induction.

This is Manber Problem 2.5.

We need to prove by induction since we don't know that the guessed form is correct. All that we **know** without doing the proof is that the form we guessed models some low-order points on the equation properly.

Technique 2: Shifted Sums

Given a sum of many terms, shift and subtract to eliminate intermediate terms.

$$G(n) = \sum_{i=0}^{n} ar^i = a + ar + ar^2 + \cdots + ar^n$$

Shift by multiplying by r.

$$rG(n) = ar + ar^2 + \cdots + ar^n + ar^{n+1}$$

Subtract.

$$G(n) - rG(n) = G(n)(1 - r) = a - ar^{n+1}$$
$$G(n) = \frac{a - ar^{n+1}}{1 - r} \quad r \neq 1$$

Technique 2: Shifted Sums
Given a sum of many terms, shift and subtract to eliminate intermediate terms.
$G(n) = \sum_{i=0}^{n} ar^i = a + ar + ar^2 + \cdots + ar^n$
Shift by multiplying by r.
$rG(n) = ar + ar^2 + \cdots + ar^n + ar^{n+1}$
Subtract.
$G(n) - rG(n) = G(n)(1 - r) = a - ar^{n+1}$
$G(n) = \frac{a - ar^{n+1}}{1 - r} \quad r \neq 1$

We often solve summations in this way – by multiplying by something or subtracting something. The big problem is that it can be a bit like finding a needle in a haystack to decide what "move" to make. We need to do something that gives us a new sum that allows us either to cancel all but a constant number of terms, or else converts all the terms into something that forms an easier summation.

Shift by multiplying by r is a reasonable guess in this example since the terms differ by a factor of r.

Example 3.3

$$G(n) = \sum_{i=1}^{n} i2^i = 1 \times 2 + 2 \times 2^2 + 3 \times 2^3 + \cdots + n \times 2^n$$

Multiply by 2.

$$2G(n) = 1 \times 2^2 + 2 \times 2^3 + 3 \times 2^4 + \cdots + n \times 2^{n+1}$$

Subtract (Note: $\sum_{i=1}^{n} 2^i = 2^{n+1} - 2$)

$$\begin{aligned} 2G(n) - G(n) &= n2^{n+1} - 2^n \cdots 2^2 - 2 \\ G(n) &= n2^{n+1} - 2^{n+1} + 2 \\ &= (n-1)2^{n+1} + 2 \end{aligned}$$

Example 3.3
$G(n) = \sum_{i=1}^{n} i2^i = 1 \times 2 + 2 \times 2^2 + 3 \times 2^3 + \cdots + n \times 2^n$
Multiply by 2.
$2G(n) = 1 \times 2^2 + 2 \times 2^3 + 3 \times 2^4 + \cdots + n \times 2^{n+1}$
Subtract (Note: $\sum_{i=1}^{n} 2^i = 2^{n+1} - 2$)
$2G(n) - G(n) = n2^{n+1} - 2^n \cdots 2^2 - 2$
$G(n) = n2^{n+1} - 2^{n+1} + 2$
$= (n-1)2^{n+1} + 2$

no notes

Recurrence Relations

- A (math) function defined in terms of itself.
- Example: Fibonacci numbers:
  $F(n) = F(n-1) + F(n-2)$ general case
  $F(1) = F(2) = 1$ base cases
- There are always one or more general cases and one or more base cases.
- We will use recurrences for time complexity of recursive (computer) functions.
- General format is $T(n) = E(T, n)$ where $E(T, n)$ is an expression in T and n.
  - $T(n) = 2T(n/2) + n$
- Alternately, an upper bound: $T(n) \leq E(T, n)$.

Recurrence Relations
A (math) function defined in terms of itself.
Example: Fibonacci numbers:
$F(n) = F(n-1) + F(n-2)$ general case
$F(1) = F(2) = 1$ base cases
There are always one or more general cases and one or more base cases.
We will use recurrences for time complexity of recursive (computer) functions.
General format is $T(n) = E(T, n)$ where $E(T, n)$ is an expression in T and n.
$T(n) = 2T(n/2) + n$
Alternately, an upper bound: $T(n) \leq E(T, n)$.

We won't spend a lot of time on techniques... just enough to be able to use them.

Solving Recurrences

We would like to find a closed form solution for $T(n)$ such that:

$$T(n) = \Theta(f(n))$$

Alternatively, find lower bound

- Not possible for inequalities of form $T(n) \leq E(T, n)$.

Methods:

- Guess (and test) a solution
- Expand recurrence
- Theorems

Solving Recurrences
We would like to find a closed form solution for $T(n)$ such that:
$T(n) = \Theta(f(n))$
Alternatively, find lower bound
Not possible for inequalities of form $T(n) \leq E(T, n)$.
Methods:
Guess (and test) a solution
Expand recurrence
Theorems

Note that "finding a closed form" means that we have $f(n)$ that doesn't include T.

Can't find lower bound for the inequality because you do not know enough... you don't know *how much bigger* $E(T, n)$ is than $T(n)$, so the result might not be $\Omega(T(n))$.

Guessing is useful for finding an asymptotic solution. Use induction to prove the guess correct.

Guessing

$T(n) = 2T(n/2) + 5n^2 \quad n \geq 2$
$T(1) = 7$
Note that T is defined only for powers of 2.

Guess a solution: $T(n) \leq c_1 n^3 = f(n)$
$T(1) = 7$ implies that $c_1 \geq 7$

Inductively, assume $T(n/2) \leq f(n/2)$.

$$\begin{aligned} T(n) &= 2T(n/2) + 5n^2 \\ &\leq 2c_1(n/2)^3 + 5n^2 \\ &\leq c_1(n^3/4) + 5n^2 \\ &\leq c_1 n^3 \text{ if } c_1 \geq 20/3. \end{aligned}$$

Guessing
$T(n) = 2T(n/2) + 5n^2 \quad n \geq 2$
$T(1) = 7$
Note that T is defined only for powers of 2.
Guess a solution: $T(n) \leq c_1 n^3 = f(n)$
$T(1) = 7$ implies that $c_1 \geq 7$.
Inductively, assume $T(n/2) \leq f(n/2)$.
$T(n) = 2T(n/2) + 5n^2$
$\leq 2c_1(n/2)^3 + 5n^2$
$\leq c_1(n^3/4) + 5n^2$
$\leq c_1 n^3$ if $c_1 \geq 20/3$.

For Big-oh, not many choices in what to guess.

$7 \times 1^3 = 7$

Because $\frac{20}{4 \cdot 3} n^3 + 5n^2 = \frac{20}{3} n^3$ when $n = 1$, and as n grows, the right side grows even faster.

Guessing (cont)

Therefore, if $c_1 = 7$, a proof by induction yields:
$T(n) \leq 7n^3$
$T(n) \in O(n^3)$

Is this the best possible solution?

Guessing (cont)
Therefore, if $c_1 = 7$, a proof by induction yields:
$T(n) \leq 7n^3$
$T(n) \in O(n^3)$
Is this the best possible solution?

No - try something tighter.

Guessing (cont)

Guess again.

$$T(n) \leq c_2 n^2 = g(n)$$

$T(1) = 7$ implies $c_2 \geq 7$.

Inductively, assume $T(n/2) \leq g(n/2)$.

$$\begin{aligned} T(n) &= 2T(n/2) + 5n^2 \\ &\leq 2c_2(n/2)^2 + 5n^2 \\ &= c_2(n^2/2) + 5n^2 \\ &\leq c_2 n^2 \text{ if } c_2 \geq 10 \end{aligned}$$

Therefore, if $c_2 = 10$, $\quad T(n) \leq 10n^2$. $\quad T(n) = O(n^2)$.

Is this the best possible upper bound?

Guessing (cont)
Guess again.
$T(n) \leq c_2 n^2 = g(n)$
$T(1) = 7$ implies $c_2 \geq 7$.
Inductively, assume $T(n/2) \leq g(n/2)$.
$T(n) = 2T(n/2) + 5n^2$
$\leq 2c_2(n/2)^2 + 5n^2$
$= c_2(n^2/2) + 5n^2$
$\leq c_2 n^2$ if $c_2 \geq 10$
Therefore, if $c_2 = 10$, $T(n) \leq 10n^2$. $T(n) = O(n^2)$.
Is this the best possible upper bound?

Because $\frac{10}{2} n^2 + 5n^2 = 10n^2$ for $n = 1$, and the right hand side grows faster.

Yes this is best, since $T(n)$ can be as bad as $5n^2$.

Guessing (cont)

Now, reshape the recurrence so that T is defined for all values of n.
$T(n) \leq 2T(\lfloor n/2 \rfloor) + 5n^2 \qquad n \geq 2$

For arbitrary n, let $2^{k-1} < n \leq 2^k$.

We have already shown that $T(2^k) \leq 10(2^k)^2$.

$$\begin{aligned} T(n) &\leq T(2^k) \leq 10(2^k)^2 \\ &= 10(2^k/n)^2 n^2 \leq 10(2)^2 n^2 \\ &\leq 40n^2 \end{aligned}$$

Hence, $T(n) = \mathrm{O}(n^2)$ for all values of n.

Typically, the bound for powers of two generalizes to all n.

no notes

Expanding Recurrences

Usually, start with equality version of recurrence.

$$\begin{aligned} T(n) &= 2T(n/2) + 5n^2 \\ T(1) &= 7 \end{aligned}$$

Assume n is a power of 2; $n = 2^k$.

no notes

Expanding Recurrences (cont)

$$\begin{aligned} T(n) &= 2T(n/2) + 5n^2 \\ &= 2(2T(n/4) + 5(n/2)^2) + 5n^2 \\ &= 2(2(2T(n/8) + 5(n/4)^2) + 5(n/2)^2) + 5n^2 \\ &= 2^k T(1) + 2^{k-1} \cdot 5(n/2^{k-1})^2 + 2^{k-2} \cdot 5(n/2^{k-2})^2 \\ &\qquad + \cdots + 2 \cdot 5(n/2)^2 + 5n^2 \\ &= 7n + 5\sum_{i=0}^{k-1} n^2/2^i = 7n + 5n^2 \sum_{i=0}^{k-1} 1/2^i \\ &= 7n + 5n^2(2 - 1/2^{k-1}) \\ &= 7n + 5n^2(2 - 2/n). \end{aligned}$$

This it the **exact** solution for powers of 2. $T(n) = \Theta(n^2)$.

no notes

Divide and Conquer Recurrences

These have the form:

$$\begin{aligned} T(n) &= aT(n/b) + cn^k \\ T(1) &= c \end{aligned}$$

... where a, b, c, k are constants.

A problem of size n is divided into a subproblems of size n/b, while cn^k is the amount of work needed to combine the solutions.

no notes

Divide and Conquer Recurrences (cont)

Expand the sum; $n = b^m$.

$$\begin{aligned} T(n) &= a(aT(n/b^2) + c(n/b)^k) + cn^k \\ &= a^m T(1) + a^{m-1}c(n/b^{m-1})^k + \cdots + ac(n/b)^k + cn^k \\ &= ca^m \sum_{i=0}^{m} (b^k/a)^i \end{aligned}$$

$a^m = a^{\log_b n} = n^{\log_b a}$
The summation is a geometric series whose sum depends on the ratio

$$r = b^k/a.$$

There are 3 cases.

Divide and Conquer Recurrences (cont)

Expand the sum; $n = b^m$.

$$\begin{aligned} T(n) &= a(aT(n/b^2) + c(n/b)^k) + cn^k \\ &= a^m T(1) + a^{m-1}c(n/b^{m-1})^k + \cdots + ac(n/b)^k + cn^k \\ &= ca^m \sum_{i=0}^{m} (b^k/a)^i \end{aligned}$$

$a^m = a^{\log_b n} = n^{\log_b a}$
The summation is a geometric series whose sum depends on the ratio

$$r = b^k/a.$$

There are 3 cases.

$n = b^m \Rightarrow m = log_b n$.

Set $a = b^{\log_b a}$. Switch order of logs, giving $(b^{\log_b n})^{\log_b a} = n^{\log_b a}$.

D & C Recurrences (cont)

(1) $r < 1$

$$\sum_{i=0}^{m} r^i < 1/(1-r), \qquad \text{a constant.}$$

$$T(n) = \Theta(a^m) = \Theta(n^{\log_b a}).$$

(2) $r = 1$

$$\sum_{i=0}^{m} r^i = m + 1 = \log_b n + 1$$

$$T(n) = \Theta(n^{\log_b a} \log n) = \Theta(n^k \log n)$$

D & C Recurrences (cont)

(1) $r < 1$

$$\sum_{i=0}^{m} r^i < 1/(1-r), \qquad \text{a constant.}$$

$$T(n) = \Theta(a^m) = \Theta(n^{\log_b a}).$$

(2) $r = 1$

$$\sum_{i=0}^{m} r^i = m + 1 = \log_b n + 1$$

$$T(n) = \Theta(n^{\log_b a} \log n) = \Theta(n^k \log n)$$

When $r = 1$, since $r = b^k/a = 1$, we get $a = b^k$.
Recall that $k = log_b a$.

D & C Recurrences (Case 3)

(3) $r > 1$

$$\sum_{i=0}^{m} r^i = \frac{r^{m+1} - 1}{r - 1} = \Theta(r^m)$$

So, from $T(n) = ca^m \sum r^i$,

$$\begin{aligned} T(n) &= \Theta(a^m r^m) \\ &= \Theta(a^m (b^k/a)^m) \\ &= \Theta(b^{km}) \\ &= \Theta(n^k) \end{aligned}$$

D & C Recurrences (Case 3)

(3) $r > 1$

$$\sum_{i=0}^{m} r^i = \frac{r^{m+1} - 1}{r - 1} = \Theta(r^m)$$

So, from $T(n) = ca^m \sum r^i$,

$$\begin{aligned} T(n) &= \Theta(a^m r^m) \\ &= \Theta(a^m (b^k/a)^m) \\ &= \Theta(b^{km}) \\ &= \Theta(n^k) \end{aligned}$$

no notes

Summary

Theorem 3.4:

$$T(n) = \begin{cases} \Theta(n^{\log_b a}) & \text{if a} > \text{b}^\text{k} \\ \Theta(n^k \log n) & \text{if a} = \text{b}^\text{k} \\ \Theta(n^k) & \text{if a} < \text{b}^\text{k} \end{cases}$$

Apply the theorem:
$T(n) = 3T(n/5) + 8n^2$.
$a = 3, b = 5, c = 8, k = 2$.
$b^k/a = 25/3$.

Case (3) holds: $T(n) = \Theta(n^2)$.

Summary

Theorem 3.4:

$$T(n) = \begin{cases} \Theta(n^{\log_b a}) & \text{if a} > \text{b}^\text{k} \\ \Theta(n^k \log n) & \text{if a} = \text{b}^\text{k} \\ \Theta(n^k) & \text{if a} < \text{b}^\text{k} \end{cases}$$

Apply the theorem:
$T(n) = 3T(n/5) + 8n^2$.
$a = 3, b = 5, c = 8, k = 2$.
$b^k/a = 25/3$.

Case (3) holds: $T(n) = \Theta(n^2)$.

We simplify by approximating summations.

Amortized Analysis

Consider this variation on STACK:

```
void init(STACK S);
element examineTop(STACK S);
void push(element x, STACK S);
void pop(int k, STACK S);
```

... where pop removes k entries from the stack.

"Local" worst case analysis for pop:
$\mathrm{O}(n)$ for n elements on the stack.

Given m_1 calls to push, m_2 calls to pop:
Naive worst case: $m_1 + m_2 \cdot n = m_1 + m_2 \cdot m_1$.

Amortized Analysis
Consider this variation on STACK:
void init(STACK S);
element examineTop(STACK S);
void push(element x, STACK S);
void pop(int k, STACK S);
... where pop removes k entries from the stack.
"Local" worst case analysis for pop:
O(n) for n elements on the stack.
Given m1 calls to push, m2 calls to pop:
Naive worst case: m1 + m2 · n = m1 + m2 · m1.

no notes

Alternate Analysis

Use amortized analysis on multiple calls to push, pop:

Cannot pop more elements than get pushed onto the stack.

After many pushes, a single pop has high **potential**.

Once that potential has been expended, it is not available for future pop operations.

The cost for m_1 pushes and m_2 pops:

$$m_1 + (m_2 + m_1) = O(m_1 + m_2)$$

Alternate Analysis
Use amortized analysis on multiple calls to push, pop:
Cannot pop more elements than get pushed onto the stack.
After many pushes, a single pop has high potential.
Once that potential has been expended, it is not available for future pop operations.
The cost for m1 pushes and m2 pops:
m1 + (m2 + m1) = O(m1 + m2)

Actual number of (constant time) push calls $+$ (Actual number of pop calls $+$ Total potential for the pops)

CLR has an entire chapter on this – we won't go into this much, but we use Amortized Analysis implicitly sometimes.