Reductions, Recursion, and Divide and Conquer

Lecture 10

Tuesday, October 1, 2024

LaTeXed: October 8, 2024 20:53

10.1
Brief intro to the RAM model

Algorithms and Computing

1. Algorithm solves a specific problem.
2. Steps/instructions of an algorithm are simple/primitive and can be executed mechanically.
3. Algorithm has a finite description; same description for all instances of the problem
4. Algorithm implicitly may have state/memory

A computer is a device that

1. implements the primitive instructions
2. allows for an automated implementation of the entire algorithm by keeping track of state

Models of Computation vs Computers

1. Model of Computation: an idealized mathematical construct that describes the primitive instructions and other details
2. Computer: an actual physical device that implements a very specific model of computation

In this course: design algorithms in a high-level model of computation.

Question: What model of computation will we use to design algorithms?

The standard programming model that you are used to in programming languages such as Java/C++. We have already seen the Turing Machine model.

Models of Computation vs Computers

1. Model of Computation: an idealized mathematical construct that describes the primitive instructions and other details
2. Computer: an actual physical device that implements a very specific model of computation

In this course: design algorithms in a high-level model of computation.

Question: What model of computation will we use to design algorithms?

The standard programming model that you are used to in programming languages such as Java/C++. We have already seen the Turing Machine model.

Unit-Cost RAM Model

Informal description:

1. Basic data type is an integer number
2. Numbers in input fit in a word
3. Arithmetic/comparison operations on words take constant time
4. Arrays allow random access (constant time to access $A[i]$)
5. Pointer based data structures via storing addresses in a word

Example

Sorting: input is an array of n numbers

1. input size is n (ignore the bits in each number),
2. comparing two numbers takes $O(1)$ time,
3. random access to array elements,
4. addition of indices takes constant time,
5. basic arithmetic operations take constant time,
6. reading/writing one word from/to memory takes constant time.

We will usually not allow (or be careful about allowing):

1. bitwise operations (and, or, xor, shift, etc).
2. floor function.
3. limit word size (usually assume unbounded word size).

Caveats of RAM Model

Unit-Cost RAM model is applicable in wide variety of settings in practice. However it is not a proper model in several important situations so one has to be careful.

1. For some problems such as basic arithmetic computation, unit-cost model makes no sense. Examples: multiplication of two n-digit numbers, primality etc.
2. Input data is very large and does not satisfy the assumptions that individual numbers fit into a word or that total memory is bounded by 2^k where k is word length.
3. Assumptions valid only for certain type of algorithms that do not create large numbers from initial data. For example, exponentiation creates very big numbers from initial numbers.

Models used in class

In this course when we design algorithms:

1. Assume unit-cost RAM by default.
2. We will explicitly point out where unit-cost RAM is not applicable for the problem at hand.
3. Turing Machines (or some high-level version of it) will be the non-cheating model that we will fall back upon when tricky issues come up.

10.1.1
What is an algorithmic problem?

What is an algorithmic problem?

Simplest and robust definition: An algorithmic problem is simply to compute a function $f : \Sigma^* \rightarrow \Sigma^*$ over strings of a finite alphabet.

Algorithm $\mathcal{A}$ solves f if for all **input strings** w, $\mathcal{A}$ outputs $f(w)$.

Typically we are interested in functions $f : D \rightarrow R$ where $D \subseteq \Sigma^*$ is the domain of f and where $R \subseteq \Sigma^*$ is the range of f.

We say that $w \in D$ is an **instance** of the problem. Implicit assumption is that the algorithm, given an arbitrary string w, can tell whether $w \in D$ or not. Parsing problem! The **size of the input** w is simply the length $|w|$.

The domain D depends on what **representation** is used. Can be lead to formally different algorithmic problems.

Types of Problems

We will broadly see three types of problems.

1. Decision Problem: Is the input a YES or NO input?
   Example: Given graph G, nodes s, t, is there a path from s to t in G?
   Example: Given a CFG grammar G and string w, is $w \in L(G)$?
2. Search Problem: Find a solution if input is a YES input.
   Example: Given graph G, nodes s, t, find an s-t path.
3. Optimization Problem: Find a best solution among all solutions for the input.
   Example: Given graph G, nodes s, t, find a shortest s-t path.

Analysis of Algorithms

Given a problem P and an algorithm $\mathcal{A}$ for P we want to know:

- Does $\mathcal{A}$ **correctly** solve problem P?
- What is the **asymptotic worst-case running time** of $\mathcal{A}$?
- What is the **asymptotic worst-case space** used by $\mathcal{A}$.

Asymptotic running-time analysis: $\mathcal{A}$ runs in $O(f(n))$ time if:

"for all n and for all inputs I of size n, $\mathcal{A}$ on input I terminates after $O(f(n))$ primitive steps."

Algorithmic Techniques

- Reduction to known problem/algorithm
- Recursion, divide-and-conquer, dynamic programming
- Graph algorithms to use as basic reductions
- Greedy

Some advanced techniques not covered in this class:

- Combinatorial optimization
- Linear and Convex Programming, more generally continuous optimization method
- Advanced data structure
- Randomization
- Many specialized areas

10.2
What is a good algorithm, and why use asymptotic running time?

What is a good algorithm?

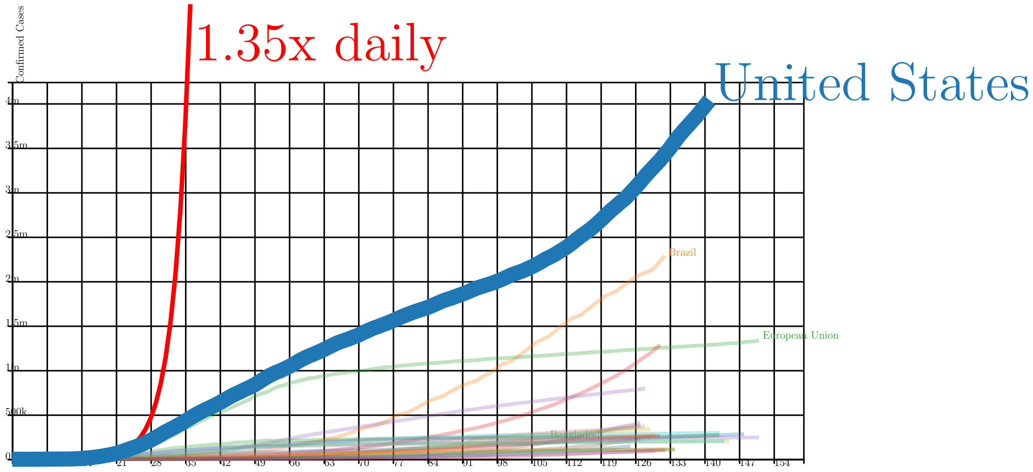
Running time...

Input size	n^2 ops	n^3 ops	n^4 ops	$n!$ ops
5	0 secs	0 secs	0 secs	0 secs
20	0 secs	0 secs	0 secs	16 mins
30	0 secs	0 secs	0 secs	$3 \cdot 10^9$ years
100	0 secs	0 secs	0 secs	never
8000	0 secs	0 secs	1 secs	never
16000	0 secs	0 secs	26 secs	never
32000	0 secs	0 secs	6 mins	never
64000	0 secs	0 secs	111 mins	never
200,000	0 secs	3 secs	7 days	never
2,000,000	0 secs	53 mins	202.943 years	never
10^8	4 secs	12.6839 years	10^9 years	never
10^9	6 mins	12683.9 years	10^{13} years	never

"No, Thursday's out. How about never—is never good for you?"

Exponential growth is bad

COVID 19 cases

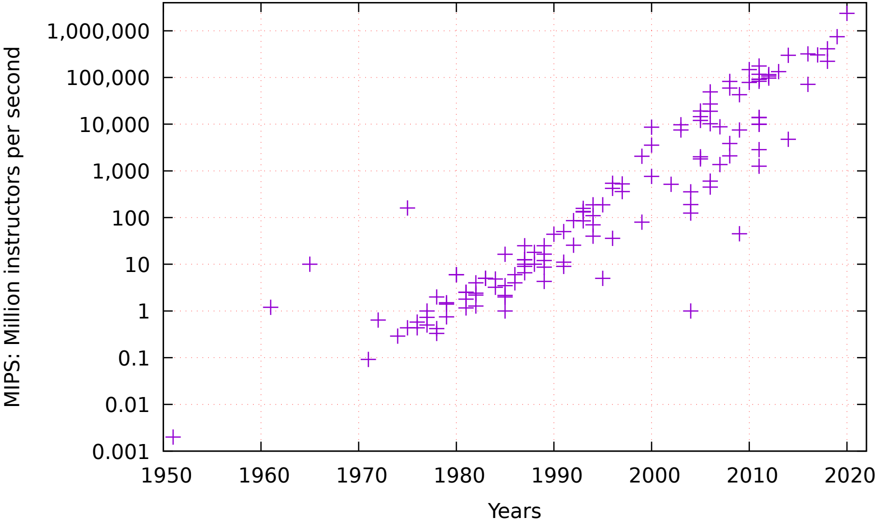


Snapshot: July 24, 2020.

CPU/Computer performance in MIPS over the years

No, no, no, exponential growth is good

https://en.wikipedia.org/wiki/Instructions_per_second



10.3
Reductions

Reduction

Reducing problem $\boldsymbol{A}$ to problem $\boldsymbol{B}$:

1. Algorithm for $\boldsymbol{A}$ uses algorithm for $\boldsymbol{B}$ as a black box

Reduction

Reducing problem A to problem B:

1. Algorithm for A uses algorithm for B as a black box

Q: How do you hunt a blue elephant?

A: With a blue elephant gun.

Reduction

Reducing problem ***A*** to problem ***B***:

1. Algorithm for ***A*** uses algorithm for ***B*** as a black box

Q: How do you hunt a blue elephant?

A: With a blue elephant gun.

Q: How do you hunt a red elephant?

A: Hold his trunk shut until it turns blue, and then shoot it with the blue elephant gun.

Reduction

Reducing problem $\boldsymbol{A}$ to problem $\boldsymbol{B}$:

1. Algorithm for $\boldsymbol{A}$ uses algorithm for $\boldsymbol{B}$ as a black box

Q: How do you hunt a blue elephant?

A: With a blue elephant gun.

Q: How do you hunt a red elephant?

A: Hold his trunk shut until it turns blue, and then shoot it with the blue elephant gun.

Q: How do you shoot a white elephant?

A: Embarrass it till it becomes red. Now use your algorithm for hunting red elephants.

UNIQUENESS: Distinct Elements Problem

Problem Given an array A of n integers, are there any duplicates in A?

Naive algorithm:

```
DistinctElements(A[1..n])
    for i = 1 to n − 1 do
        for j = i + 1 to n do
            if (A[i] = A[j])
                return YES
    return NO
```

Running time: $O(n^2)$

UNIQUENESS: Distinct Elements Problem

Problem Given an array A of n integers, are there any duplicates in A?

Naive algorithm:

```
DistinctElements(A[1..n])
    for i = 1 to n − 1 do
        for j = i + 1 to n do
            if (A[i] = A[j])
                return YES
    return NO
```

Running time: $O(n^2)$

UNIQUENESS: Distinct Elements Problem

Problem Given an array A of n integers, are there any duplicates in A?

Naive algorithm:

```
DistinctElements(A[1..n])
    for i = 1 to n − 1 do
        for j = i + 1 to n do
            if (A[i] = A[j])
                return YES
    return NO
```

Running time: $O(n^2)$

UNIQUENESS: Distinct Elements Problem

Problem Given an array A of n integers, are there any duplicates in A?

Naive algorithm:

```
DistinctElements(A[1..n])
    for i = 1 to n − 1 do
        for j = i + 1 to n do
            if (A[i] = A[j])
                return YES
    return NO
```

Running time: $O(n^2)$

Reduction to Sorting

```
DistinctElements(A[1..n])
    Sort A
    for i = 1 to n − 1 do
        if (A[i] = A[i + 1]) then
            return YES
    return NO
```

Running time: $O(n)$ plus time to sort an array of n numbers

Important point: algorithm uses sorting as a black box

Advantage of naive algorithm: works for objects that cannot be "sorted". Can also consider hashing but outside scope of current course.

Reduction to Sorting

```
DistinctElements(A[1..n])
    Sort A
    for i = 1 to n − 1 do
        if (A[i] = A[i + 1]) then
            return YES
    return NO
```

Running time: $O(n)$ plus time to sort an array of n numbers

Important point: algorithm uses sorting as a black box

Advantage of naive algorithm: works for objects that cannot be "sorted". Can also consider hashing but outside scope of current course.

Reduction to Sorting

```
DistinctElements(A[1..n])
    Sort A
    for i = 1 to n − 1 do
        if (A[i] = A[i + 1]) then
            return YES
    return NO
```

Running time: $O(n)$ plus time to sort an array of n numbers

Important point: algorithm uses sorting as a black box

Advantage of naive algorithm: works for objects that cannot be "sorted". Can also consider hashing but outside scope of current course.

Two sides of Reductions

Suppose problem A reduces to problem B

1. Positive direction: Algorithm for B implies an algorithm for A
2. Negative direction: Suppose there is no "efficient" algorithm for A then it implies no efficient algorithm for B (technical condition for reduction time necessary for this)

Example: Distinct Elements reduces to Sorting in $O(n)$ time

1. An $O(n \log n)$ time algorithm for Sorting implies an $O(n \log n)$ time algorithm for Distinct Elements problem.
2. If there is no $o(n \log n)$ time algorithm for Distinct Elements problem then there is no $o(n \log n)$ time algorithm for Sorting.

Two sides of Reductions

Suppose problem A reduces to problem B

1. Positive direction: Algorithm for B implies an algorithm for A
2. Negative direction: Suppose there is no "efficient" algorithm for A then it implies no efficient algorithm for B (technical condition for reduction time necessary for this)

Example: Distinct Elements reduces to Sorting in $O(n)$ time

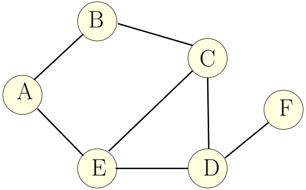
1. An $O(n \log n)$ time algorithm for Sorting implies an $O(n \log n)$ time algorithm for Distinct Elements problem.
2. If there is no $o(n \log n)$ time algorithm for Distinct Elements problem then there is no $o(n \log n)$ time algorithm for Sorting.

10.3.1
More examples of reductions

Maximum Independent Set in a Graph

Definition 10.1.

Given undirected graph $G = (V, E)$ a subset of nodes $S \subseteq V$ is an independent set (also called a stable set) if for there are no edges between nodes in S. That is, if $u, v \in S$ then $(u, v) \notin E$.

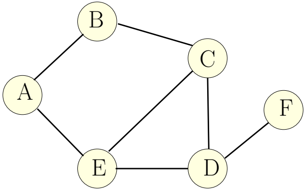


Some independent sets in graph above:

Maximum Independent Set Problem

Input Graph $G = (V, E)$

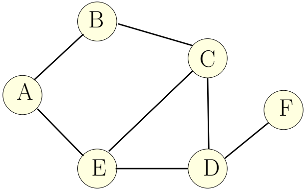
Goal Find maximum sized independent set in G



Maximum Weight Independent Set Problem

Input Graph $G = (V, E)$, weights $w(v) \geq 0$ for $v \in V$

Goal Find maximum weight independent set in G

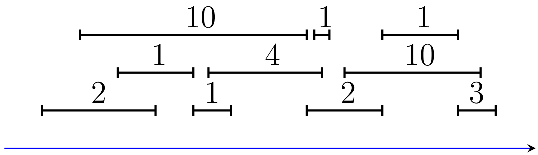


Weighted Interval Scheduling

Input A set of jobs with start times, finish times and weights (or profits).

Goal Schedule jobs so that total weight of jobs is maximized.

1. Two jobs with overlapping intervals cannot both be scheduled!

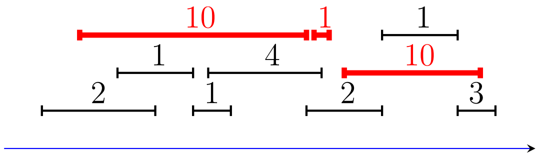


Weighted Interval Scheduling

Input A set of jobs with start times, finish times and weights (or profits).

Goal Schedule jobs so that total weight of jobs is maximized.

1. Two jobs with overlapping intervals cannot both be scheduled!



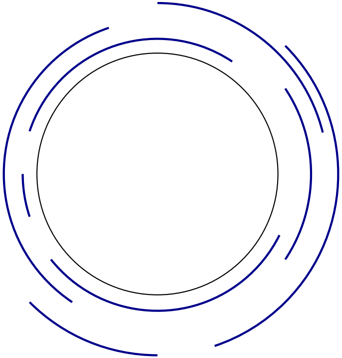
Reduction from Interval Scheduling to MIS

Question: Can you reduce Weighted Interval Scheduling to Max Weight Independent Set Problem?

Weighted Circular Arc Scheduling

Input A set of arcs on a circle, each arc has a weight (or profit).

Goal Find a maximum weight subset of arcs that do not overlap.



Reductions

Question: Can you reduce Weighted Interval Scheduling to Weighted Circular Arc Scheduling?

Question: Can you reduce Weighted Circular Arc Scheduling to Weighted Interval Scheduling? Yes!

```
MaxWeightIndependentArcs(arcs 𝒞)
    cur-max = 0
    for each arc C ∈ 𝒞 do
        Remove C and all arcs overlapping with C
        w_C = wt of opt. solution in resulting Interval problem
        w_C = w_C + wt(C)
        cur-max = max{cur-max, w_C}
    end for
    return cur-max
```

n calls to the sub-routine for interval scheduling

Reductions

Question: Can you reduce Weighted Interval Scheduling to Weighted Circular Arc Scheduling?

Question: Can you reduce Weighted Circular Arc Scheduling to Weighted Interval Scheduling? Yes!

```
MaxWeightIndependentArcs(arcs 𝒞)
    cur-max = 0
    for each arc C ∈ 𝒞 do
        Remove C and all arcs overlapping with C
        w_C = wt of opt. solution in resulting Interval problem
        w_C = w_C + wt(C)
        cur-max = max{cur-max, w_C}
    end for
    return cur-max
```

n calls to the sub-routine for interval scheduling

Reductions

Question: Can you reduce Weighted Interval Scheduling to Weighted Circular Arc Scheduling?

Question: Can you reduce Weighted Circular Arc Scheduling to Weighted Interval Scheduling? Yes!

```
MaxWeightIndependentArcs(arcs 𝒞)
    cur-max = 0
    for each arc C ∈ 𝒞 do
        Remove C and all arcs overlapping with C
        w_C = wt of opt. solution in resulting Interval problem
        w_C = w_C + wt(C)
        cur-max = max{cur-max, w_C}
    end for
    return cur-max
```

n calls to the sub-routine for interval scheduling

Reductions

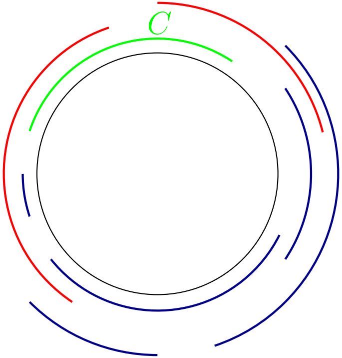
Question: Can you reduce Weighted Interval Scheduling to Weighted Circular Arc Scheduling?

Question: Can you reduce Weighted Circular Arc Scheduling to Weighted Interval Scheduling? Yes!

```
MaxWeightIndependentArcs(arcs 𝒞)
    cur-max = 0
    for each arc C ∈ 𝒞 do
        Remove C and all arcs overlapping with C
        w_C = wt of opt. solution in resulting Interval problem
        w_C = w_C + wt(C)
        cur-max = max{cur-max, w_C}
    end for
    return cur-max
```

n calls to the sub-routine for interval scheduling

Illustration



10.4
Recursion as self reductions

Recursion

Reduction: reduce one problem to another

Recursion: a special case of reduction

1. reduce problem to a smaller instance of itself
2. self-reduction

1. Problem instance of size n is reduced to one or more instances of size $n-1$ or less.
2. For termination, problem instances of small size are solved by some other method as base cases

Recursion

Reduction: reduce one problem to another

Recursion: a special case of reduction

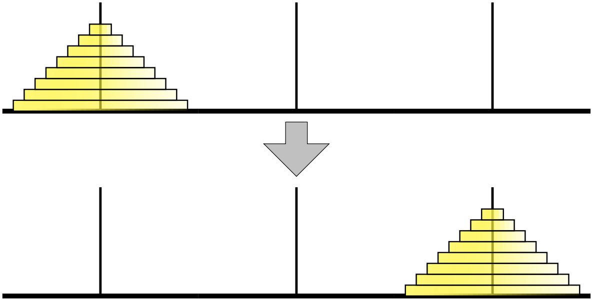
1. reduce problem to a smaller instance of itself
2. self-reduction

1. Problem instance of size n is reduced to one or more instances of size $n - 1$ or less.
2. For termination, problem instances of small size are solved by some other method as base cases

Recursion

1. Recursion is a very powerful and fundamental technique
2. Basis for several other methods
   2.1 Divide and conquer
   2.2 Dynamic programming
   2.3 Enumeration and branch and bound etc
   2.4 Some classes of greedy algorithms
3. Makes proof of correctness easy (via induction)
4. Recurrences arise in analysis

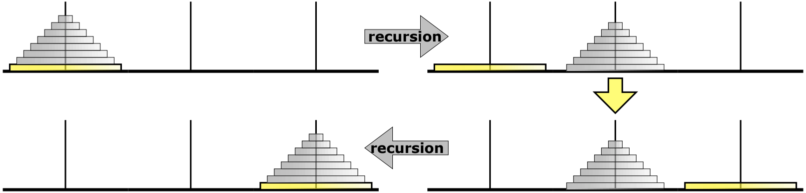
Tower of Hanoi



The Tower of Hanoi puzzle

Move stack of n disks from peg $\mathbf{0}$ to peg $\mathbf{2}$, one disk at a time.
Rule: cannot put a larger disk on a smaller disk.
Question: what is a strategy and how many moves does it take?

Tower of Hanoi via Recursion



The Tower of Hanoi algorithm; ignore everything but the bottom disk

Recursive Algorithm

```
Hanoi(n, src, dest, tmp):
    if (n > 0) then
        Hanoi(n − 1, src, tmp, dest)
        Move disk n from src to dest
        Hanoi(n − 1, tmp, dest, src)
```

$T(n)$: time to move n disks via recursive strategy

$$T(n) = 2T(n-1) + 1 \qquad n > 1 \qquad \text{and } T(1) = 1$$

Recursive Algorithm

```
Hanoi(n, src, dest, tmp):
    if (n > 0) then
        Hanoi(n − 1, src, tmp, dest)
        Move disk n from src to dest
        Hanoi(n − 1, tmp, dest, src)
```

$T(n)$: time to move n disks via recursive strategy

$$T(n) = 2T(n-1) + 1 \qquad n > 1 \qquad \text{and } T(1) = 1$$

Recursive Algorithm

```
Hanoi(n, src, dest, tmp):
    if (n > 0) then
        Hanoi(n − 1, src, tmp, dest)
        Move disk n from src to dest
        Hanoi(n − 1, tmp, dest, src)
```

$T(n)$: time to move n disks via recursive strategy

$$T(n) = 2T(n-1) + 1 \qquad n > 1 \qquad \text{and } T(1) = 1$$

Analysis

$$
\begin{aligned}
T(n) &= 2T(n-1)+1 \\
&= 2^2T(n-2)+2+1 \\
&= \ldots \\
&= 2^iT(n-i)+2^{i-1}+2^{i-2}+\ldots+1 \\
&= \ldots \\
&= 2^{n-1}T(1)+2^{n-2}+\ldots+1 \\
&= 2^{n-1}+2^{n-2}+\ldots+1 \\
&= (2^n-1)/(2-1)=2^n-1
\end{aligned}
$$

10.5
Divide and Conquer

Divide and Conquer Paradigm

Divide and Conquer is a common and useful type of recursion

Approach

1. Break problem instance into smaller instances - divide step
2. Recursively solve problem on smaller instances
3. Combine solutions to smaller instances to obtain a solution to the original instance - conquer step

Question: Why is this not plain recursion?

1. In divide and conquer, each smaller instance is typically at least a constant factor smaller than the original instance which leads to efficient running times.
2. There are many examples of this particular type of recursion that it deserves its own treatment.

Divide and Conquer Paradigm

Divide and Conquer is a common and useful type of recursion

Approach

1. Break problem instance into smaller instances - divide step
2. Recursively solve problem on smaller instances
3. Combine solutions to smaller instances to obtain a solution to the original instance - conquer step

Question: Why is this not plain recursion?

1. In divide and conquer, each smaller instance is typically at least a constant factor smaller than the original instance which leads to efficient running times.
2. There are many examples of this particular type of recursion that it deserves its own treatment.

Divide and Conquer Paradigm

Divide and Conquer is a common and useful type of recursion

Approach

1. Break problem instance into smaller instances - divide step
2. Recursively solve problem on smaller instances
3. Combine solutions to smaller instances to obtain a solution to the original instance - conquer step

Question: Why is this not plain recursion?

1. In divide and conquer, each smaller instance is typically at least a constant factor smaller than the original instance which leads to efficient running times.
2. There are many examples of this particular type of recursion that it deserves its own treatment.

10.6
Merge Sort

Sorting

Input Given an array of n elements

Goal Rearrange them in ascending order

Merge Sort [von Neumann]

MergeSort

1. Input: Array $\boldsymbol{A[1 \ldots n]}$

$$\boldsymbol{A\ L\ G\ O\ R\ I\ T\ H\ M\ S}$$

Merge Sort [von Neumann]

MergeSort

1. Input: Array $A[1 \dots n]$

$$A \; L \; G \; O \; R \; I \; T \; H \; M \; S$$

2. Divide into subarrays $A[1 \dots m]$ and $A[m+1 \dots n]$, where $m = \lfloor n/2 \rfloor$

$$A \; L \; G \; O \; R \qquad I \; T \; H \; M \; S$$

Merge Sort [von Neumann]

MergeSort

1. Input: Array $A[1 \dots n]$

   $$A\ L\ G\ O\ R\ I\ T\ H\ M\ S$$

2. Divide into subarrays $A[1 \dots m]$ and $A[m+1 \dots n]$, where $m = \lfloor n/2 \rfloor$

   $$A\ L\ G\ O\ R \qquad I\ T\ H\ M\ S$$

3. Recursively **MergeSort** $A[1 \dots m]$ and $A[m+1 \dots n]$

   $$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$

Merge Sort [von Neumann]

MergeSort

1. Input: Array $A[1 \dots n]$

   $$A\ L\ G\ O\ R\ I\ T\ H\ M\ S$$

2. Divide into subarrays $A[1 \dots m]$ and $A[m+1 \dots n]$, where $m = \lfloor n/2 \rfloor$

   $$A\ L\ G\ O\ R \qquad I\ T\ H\ M\ S$$

3. Recursively **MergeSort** $A[1 \dots m]$ and $A[m+1 \dots n]$

   $$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$

4. Merge the sorted arrays

   $$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

Merge Sort [von Neumann]

MergeSort

1. Input: Array $A[1 \dots n]$

   $A\ L\ G\ O\ R\ I\ T\ H\ M\ S$

2. Divide into subarrays $A[1 \dots m]$ and $A[m+1 \dots n]$, where $m = \lfloor n/2 \rfloor$

   $A\ L\ G\ O\ R \qquad I\ T\ H\ M\ S$

3. Recursively **MergeSort** $A[1 \dots m]$ and $A[m+1 \dots n]$

   $A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$

4. Merge the sorted arrays

   $A\ G\ H\ I\ L\ M\ O\ R\ S\ T$

Merging Sorted Arrays

1. Use a new array $\boldsymbol{C}$ to store the merged array
2. Scan $\boldsymbol{A}$ and $\boldsymbol{B}$ from left-to-right, storing elements in $\boldsymbol{C}$ in order

$$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$

$$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

Merging Sorted Arrays

1. Use a new array C to store the merged array
2. Scan A and B from left-to-right, storing elements in C in order

$$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$
$$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

Merging Sorted Arrays

1. Use a new array C to store the merged array
2. Scan A and B from left-to-right, storing elements in C in order

$$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$

$$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

Merging Sorted Arrays

1. Use a new array C to store the merged array
2. Scan A and B from left-to-right, storing elements in C in order

$$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$
$$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

Merging Sorted Arrays

1. Use a new array C to store the merged array
2. Scan A and B from left-to-right, storing elements in C in order

$$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$
$$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

Merging Sorted Arrays

1. Use a new array C to store the merged array
2. Scan A and B from left-to-right, storing elements in C in order

$$A\ G\ L\ O\ R \qquad H\ I\ M\ S\ T$$
$$A\ G\ H\ I\ L\ M\ O\ R\ S\ T$$

3. Merge two arrays using only constantly more extra space (in-place merge sort): doable but complicated and typically impractical.

Formal Code

```
MERGESORT(A[1..n]):
  if n > 1
      m ← ⌊n/2⌋
      MERGESORT(A[1..m])
      MERGESORT(A[m+1..n])
      MERGE(A[1..n], m)
```

```
MERGE(A[1..n], m):
  i ← 1;  j ← m+1
  for k ← 1 to n
      if j > n
          B[k] ← A[i];  i ← i+1
      else if i > m
          B[k] ← A[j];  j ← j+1
      else if A[i] < A[j]
          B[k] ← A[i];  i ← i+1
      else
          B[k] ← A[j];  j ← j+1

  for k ← 1 to n
      A[k] ← B[k]
```

Merge using lists

```
List MergeL(List a, List b)
    if a = NULL then return b
    if b = NULL then return a

    if a.data ≤ b.data then
        return List( a.date, MergeL( a.tail, b ) )
    else
        return List( b.data, MergeL( a, b.tail ) )
```

10.6.1
Proving that merge is correct

Proving Correctness

Obvious way to prove correctness of recursive algorithm: induction!

- Easy to show by induction on n that MergeSort is correct if you assume Merge is correct.
- How do we prove that Merge is correct? Also by induction!
- One way is to rewrite Merge into a recursive version.
- For algorithms with loops one comes up with a natural loop invariant that captures all the essential properties and then we prove the loop invariant by induction on the index of the loop.

Proving Correctness

Obvious way to prove correctness of recursive algorithm: induction!

- Easy to show by induction on n that MergeSort is correct if you assume Merge is correct.
- How do we prove that Merge is correct? Also by induction!
- One way is to rewrite Merge into a recursive version.
- For algorithms with loops one comes up with a natural loop invariant that captures all the essential properties and then we prove the loop invariant by induction on the index of the loop.

Merge is correct..

```
Merge(A[1...m], A[m+1...n])
    i ← 1, j ← m+1, k ← 1
    while ( k ≤ n ) do
        if i > m or (j ≤ n and A[i] > A[j])
            B[k++] ← A[j++]
        else
            B[k++] ← A[i++]
    A ← B
```

Claim 10.1.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct). For any value of k, in the beginning of the loop, we have:

1. *$B[1...k-1]$ contains the $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Merge is correct..

```
Merge(A[1...m], A[m+1...n])
    i ← 1, j ← m+1, k ← 1
    while ( k ≤ n ) do
        if i > m or (j ≤ n and A[i] > A[j])
            B[k++] ← A[j++]
        else
            B[k++] ← A[i++]
    A ← B
```

Claim 10.1.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct). For any value of k, in the beginning of the loop, we have:

1. $B[1...k-1]$ *contains the* $k-1$ *smallest elements in* A.
2. $B[1...k-1]$ *is sorted.*

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Proof:

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Proof:

Base of induction: $k = 1$: Emptily true.

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Proof:

Inductive hypothesis: Claim true for all $k \leq \alpha$.

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1,  j ← m+1,  k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Proof:

Inductive hypothesis: Claim true for all $k \leq \alpha$.
Inductive step: Need to prove claim true for $k = \alpha + 1$.

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Inductive hypothesis: Claim true for all $k \leq \alpha$.
Idea: Start at iteration $k = \alpha$, and use induction hypothesis, run the loop for one iter...

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Inductive hypothesis: Claim true for all $k \leq \alpha$.
Idea: Start at iteration $k = \alpha$, and use induction hypothesis, run the loop for one iter...

If $i > m$ then true.

Merge is correct

Merge$(A[1\dots m], A[m+1\dots n])$
$i \leftarrow 1,\ j \leftarrow m+1,\ k \leftarrow 1$
while ($k \leq n$) **do**
    if $i > m$ or ($j \leq n$ and $A[i] > A[j]$)
        $B[k++] \leftarrow A[j++]$
    else
        $B[k++] \leftarrow A[i++]$
$A \leftarrow B$

Claim 10.2.

Assuming $A[1\dots m]$ and $A[m+1\dots n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1\dots k-1]$: $k-1$ smallest elements in A.*
2. *$B[1\dots k-1]$ is sorted.*

Inductive hypothesis: Claim true for all $k \leq \alpha$.
Idea: Start at iteration $k = \alpha$, and use induction hypothesis, run the loop for one iter...

If $i > m$ then true.
If $j > n$ then true.

Merge is correct

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Inductive hypothesis: Claim true for all $k \leq \alpha$.
Idea: Start at iteration $k = \alpha$, and use induction hypothesis, run the loop for one iter...

If $i \leq m$ and $j \leq n$ then...

Merge is correct

```
Merge(A[1...m], A[m + 1...n])
i ← 1, j ← m + 1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

Claim 10.2.

Assuming $A[1...m]$ and $A[m+1...n]$ are sorted (all values distinct).
$\forall k$, in beginning of the loop, we have:

1. *$B[1...k-1]$: $k-1$ smallest elements in A.*
2. *$B[1...k-1]$ is sorted.*

Inductive hypothesis: Claim true for all $k \leq \alpha$.
Idea: Start at iteration $k = \alpha$, and use induction hypothesis, run the loop for one iter...

If $i \leq m$ and $j \leq n$ then...

□

Merge is correct!!!

Claim 10.3.

Assuming $A[1 \ldots m]$ *and* $A[m+1 \ldots n]$ *are sorted (all values distinct).*
$\forall k$, *in beginning of the loop, we have:*

1. $B[1 \ldots k-1]$: $k-1$ *smallest elements in* A.
2. $B[1 \ldots k-1]$ *is sorted.*

Proved claim is correct. Plugging $k = n+1$, implies.

Claim 10.4.

By end of loop execution B *(and thus* A*) contain the elements of* A *in sorted order.*
$\implies$ **Merge** *is correct.*

10.6.2
Proving that merge-sort is correct

Proving correctness of merge-sort

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m+1...n])
    Merge(A[1...m], A(m+1...n])
```

Proved: Merge is correct.

Proving correctness of merge-sort

```
Merge(A[1...m], A[m+1...n])
i ← 1, j ← m+1, k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m+1...n])
    Merge(A[1...m], A(m+1...n])
```

Proved: Merge is correct.

Lemma 10.5.

MergeSort *correctly sort the input array.*

Proving correctness of merge-sort

```
Merge(A[1...m], A[m+1...n])
i ← 1,  j ← m+1,  k ← 1
while ( k ≤ n ) do
    if i > m or (j ≤ n and A[i] > A[j])
        B[k++] ← A[j++]
    else
        B[k++] ← A[i++]
A ← B
```

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m+1...n])
    Merge(A[1...m], A(m+1...n])
```

Lemma 10.5.

MergeSort *correctly sort the input array.*

Proof by induction on n.

Proving correctness of merge-sort

Lemma 10.6.
MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m + 1...n])
    Merge(A[1...m], A(m + 1...n])
```

Proof: By induction on n.

Proving correctness of merge-sort

Lemma 10.6.
MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m + 1...n])
    Merge(A[1...m], A(m + 1...n])
```

Proof: By induction on n.
Base: $n = 1$.

Proving correctness of merge-sort

Lemma 10.6.
MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m + 1...n])
    Merge(A[1...m], A(m + 1...n])
```

Proof: By induction on n.
Base: $n = 1$.
Inductive hypothesis Lemma correct for all $n \leq k$.

Proving correctness of merge-sort

Lemma 10.6.

MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m + 1...n])
    Merge(A[1...m], A(m + 1...n])
```

Proof: By induction on n.

Base: $n = 1$.

Inductive hypothesis Lemma correct for all $n \leq k$.

Inductive step: Need to prove that lemma holds for $n = k + 1 \geq 2$.

Proving correctness of merge-sort

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m + 1...n])
    Merge(A[1...m], A(m + 1...n])
```

Lemma 10.6.

MergeSort *correctly sort the input array.*

Proof: By induction on n.

Base: $n = 1$.

Inductive hypothesis Lemma correct for all $n \leq k$.

Inductive step: Need to prove that lemma holds for $n = k + 1 \geq 2$.

$m = \lfloor n/2 \rfloor < n$: Can use induction on $A[1...m]$.

Proving correctness of merge-sort

Lemma 10.6.
MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m+1...n])
    Merge(A[1...m], A(m+1...n])
```

Proof: By induction on n.
Base: $n = 1$.
Inductive hypothesis Lemma correct for all $n \leq k$.
Inductive step: Need to prove that lemma holds for $n = k + 1 \geq 2$.
$m = \lfloor n/2 \rfloor < n$: Can use induction on $A[1...m]$.
$n - m < n$: Can use induction on $A[m + 1...n]$.

Proving correctness of merge-sort

Lemma 10.6.

MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m+1...n])
    Merge(A[1...m], A(m+1...n])
```

Proof: By induction on n.

Base: $n = 1$.

Inductive hypothesis Lemma correct for all $n \leq k$.

Inductive step: Need to prove that lemma holds for $n = k + 1 \geq 2$.

$m = \lfloor n/2 \rfloor < n$: Can use induction on $A[1...m]$.

$n - m < n$: Can use induction on $A[m + 1...n]$.

$\implies A[1...m], A[m + 1...n]$ are sorted correctly. by induction.

Proving correctness of merge-sort

Lemma 10.6.

MergeSort *correctly sort the input array.*

```
MergeSort(A[1...n])
    if n ≤ 1 then return
    m ← ⌊n/2⌋
    MergeSort(A[1...m])
    MergeSort(A[m+1...n])
    Merge(A[1...m], A(m+1...n])
```

Proof: By induction on n.

Base: $n = 1$.

Inductive hypothesis Lemma correct for all $n \leq k$.

Inductive step: Need to prove that lemma holds for $n = k + 1 \geq 2$.

$m = \lfloor n/2 \rfloor < n$: Can use induction on $A[1...m]$.

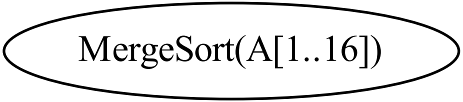
$n - m < n$: Can use induction on $A[m+1...n]$.

$\implies$ $A[1...m]$, $A[m+1...n]$ are sorted correctly. by induction.

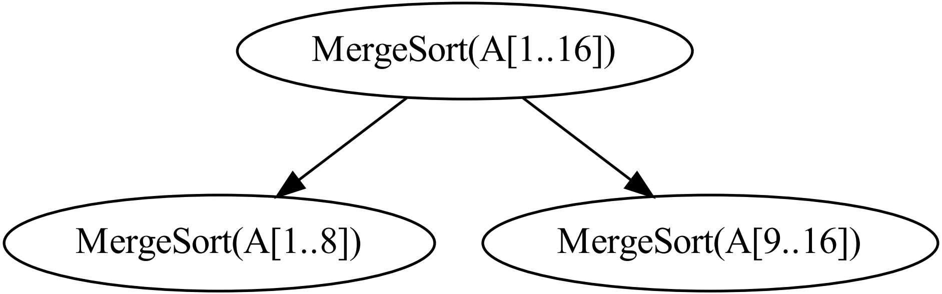
Since Merge is correct $\implies$ $A[1...n]$ is sorted correctly.

10.6.3
Running time analysis of merge-sort: Recursion tree method

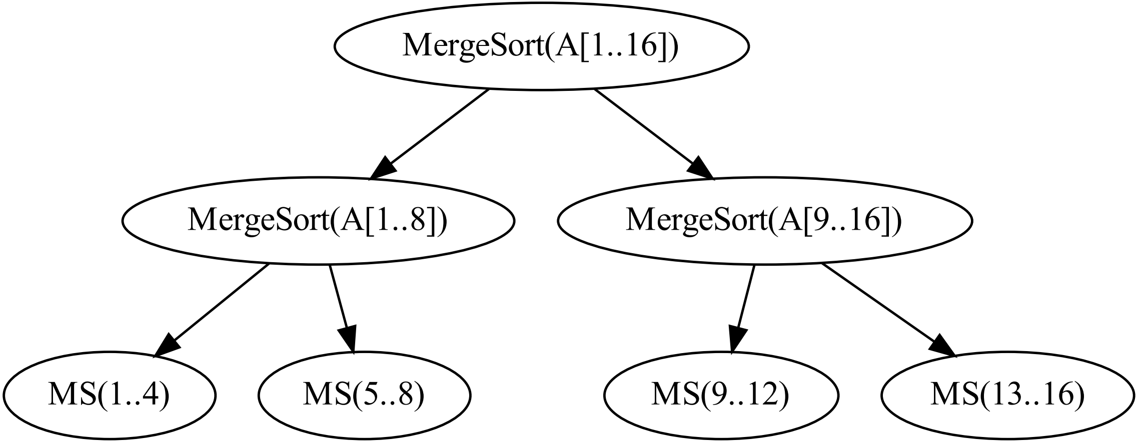
Recursion tree



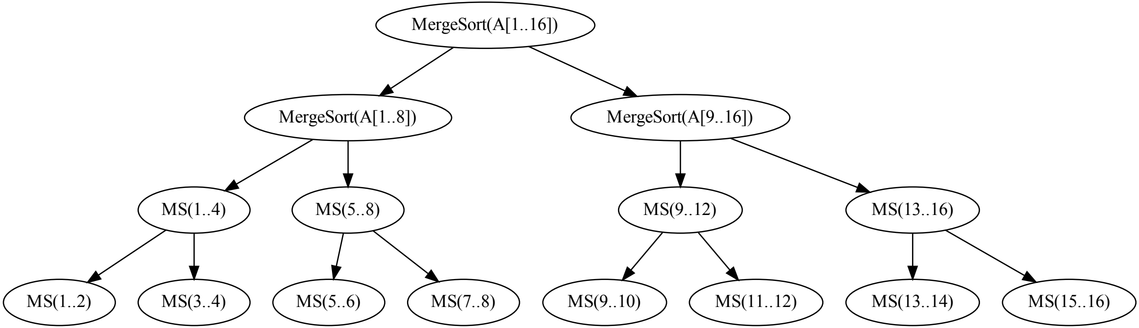
Recursion tree



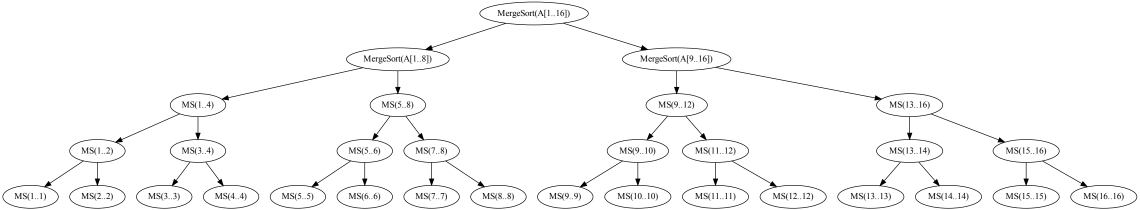
Recursion tree



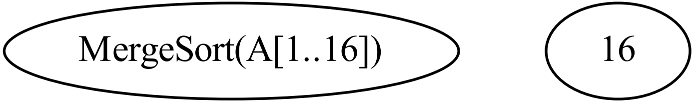
Recursion tree



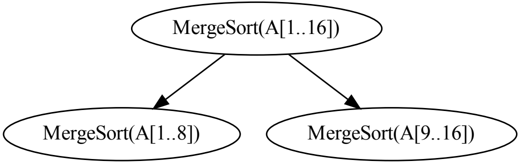
Recursion tree

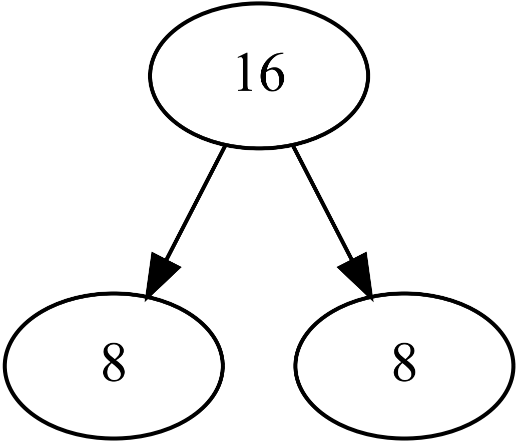


Recursion tree: subproblem sizes

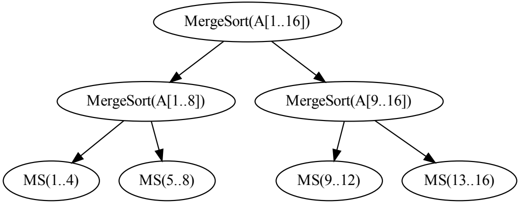


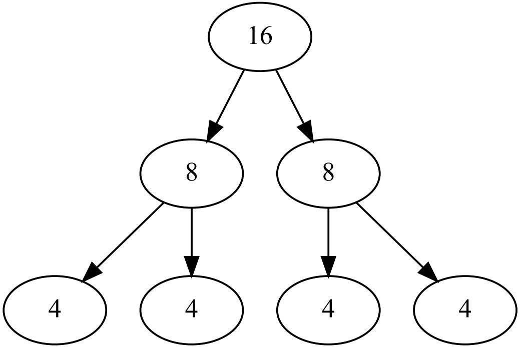
Recursion tree: subproblem sizes



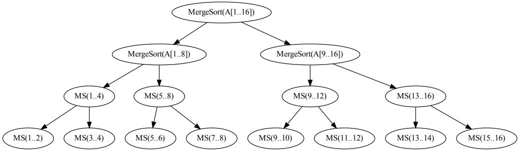


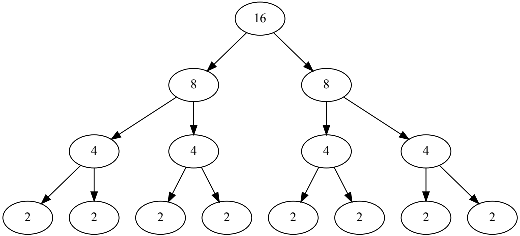
Recursion tree: subproblem sizes



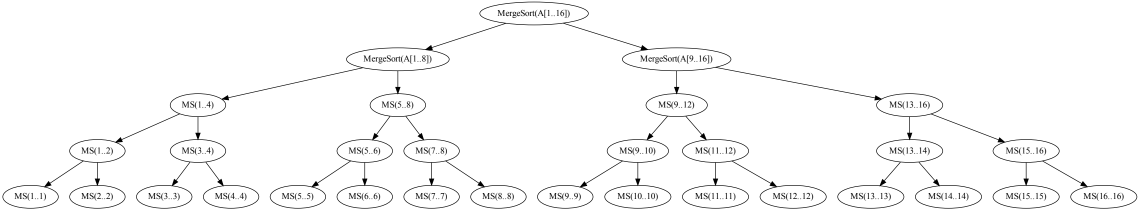


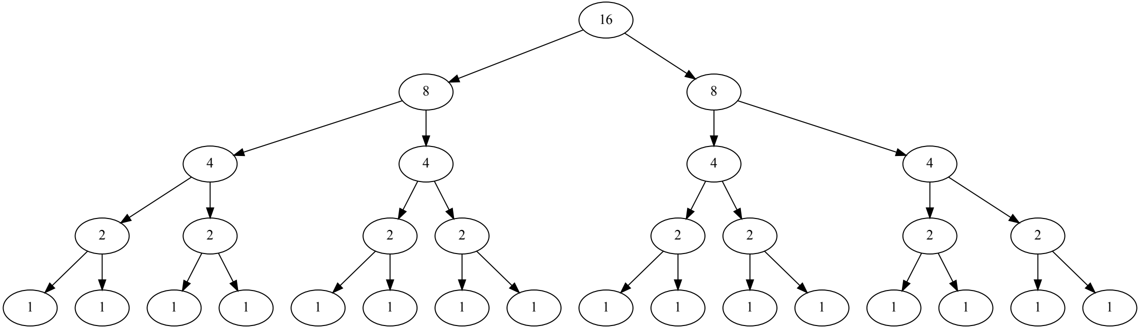
Recursion tree: subproblem sizes



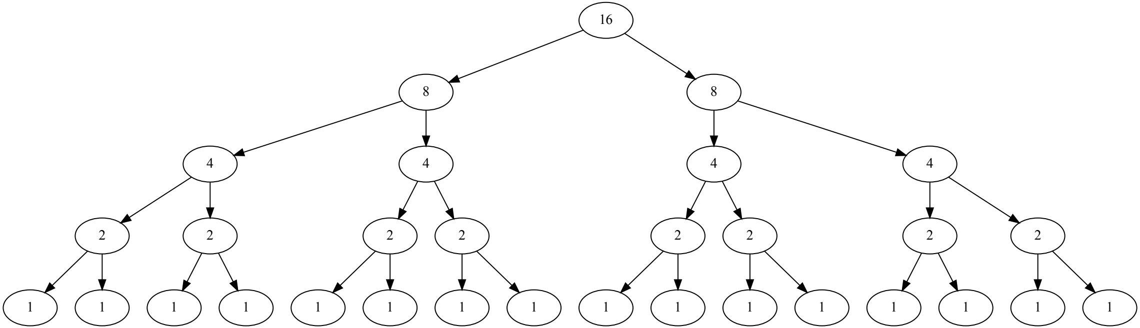


Recursion tree: subproblem sizes





Recursion tree: Total work?



Running Time

$T(n)$: time for merge sort to sort an n element array

$$T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + cn$$

What do we want as a solution to the recurrence?

Almost always only an <u>asymptotically</u> tight bound. That is we want to know $f(n)$ such that $T(n) = \Theta(f(n))$.

1. $T(n) = O(f(n))$ - upper bound
2. $T(n) = \Omega(f(n))$ - lower bound

Running Time

$T(n)$: time for merge sort to sort an n element array

$$T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + cn$$

What do we want as a solution to the recurrence?

Almost always only an asymptotically tight bound. That is we want to know $f(n)$ such that $T(n) = \Theta(f(n))$.

1. $T(n) = O(f(n))$ - upper bound
2. $T(n) = \Omega(f(n))$ - lower bound

Running Time

$T(n)$: time for merge sort to sort an n element array

$$T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + cn$$

What do we want as a solution to the recurrence?

Almost always only an <u>asymptotically</u> tight bound. That is we want to know $f(n)$ such that $T(n) = \Theta(f(n))$.

1. $T(n) = O(f(n))$ - upper bound
2. $T(n) = \Omega(f(n))$ - lower bound

Solving Recurrences: Some Techniques

1. Know some basic math: geometric series, logarithms, exponentials, elementary calculus
2. Expand the recurrence and spot a pattern and use simple math
3. Recursion tree method — imagine the computation as a tree
4. Guess and verify — useful for proving upper and lower bounds even if not tight bounds

Albert Einstein: "Everything should be made as simple as possible, but not simpler."

Know where to be loose in analysis and where to be tight. Comes with practice, practice, practice!

Review notes on recurrence solving.

Solving Recurrences: Some Techniques

1. Know some basic math: geometric series, logarithms, exponentials, elementary calculus
2. Expand the recurrence and spot a pattern and use simple math
3. Recursion tree method — imagine the computation as a tree
4. Guess and verify — useful for proving upper and lower bounds even if not tight bounds

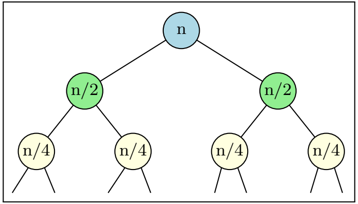
Albert Einstein: "Everything should be made as simple as possible, but not simpler."

Know where to be loose in analysis and where to be tight. Comes with practice, practice, practice!

Review notes on recurrence solving.

Recursion Trees

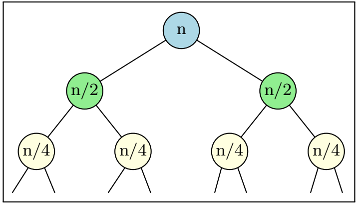
MergeSort: **n** is a power of **2**



1. Unroll the recurrence. $T(n) = 2T(n/2) + cn$

Recursion Trees

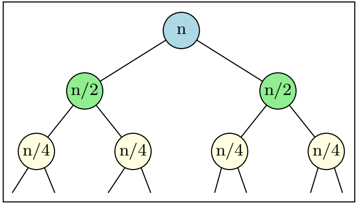
MergeSort: **n** is a power of **2**



1. Unroll the recurrence. $T(n) = 2T(n/2) + cn$
2. Identify a pattern. At the ith level total work is cn.

Recursion Trees

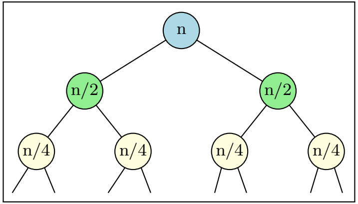
MergeSort: **n** is a power of **2**



1. Unroll the recurrence. $T(n) = 2T(n/2) + cn$
2. Identify a pattern. At the ith level total work is cn.

Recursion Trees

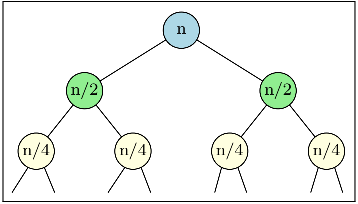
MergeSort: $\mathbf{n}$ is a power of $\mathbf{2}$



1. Unroll the recurrence. $T(n) = 2T(n/2) + cn$
2. Identify a pattern. At the ith level total work is cn.
3. Sum over all levels. The number of levels is $\log n$.
   So total is $cn \log n = O(n \log n)$.

Recursion Trees

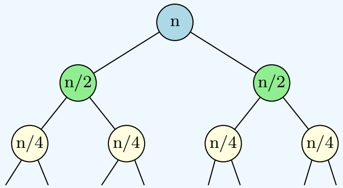
MergeSort: **n** is a power of **2**



1. Unroll the recurrence. $T(n) = 2T(n/2) + cn$
2. Identify a pattern. At the ith level total work is cn.
3. Sum over all levels. The number of levels is $\log n$.
   So total is $cn \log n = O(n \log n)$.

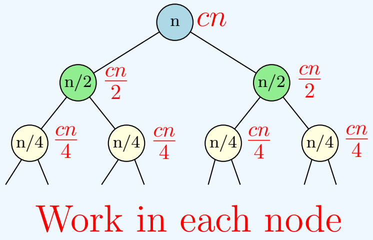
Recursion Trees

An illustrated example...



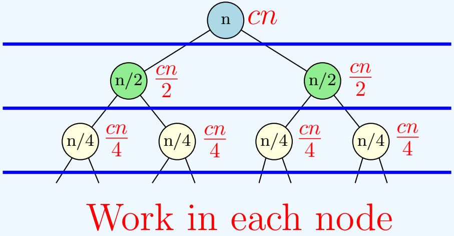
Recursion Trees

An illustrated example...



Recursion Trees

An illustrated example...



Recursion Trees

An illustrated example...

$\log n$ levels:

$$cn = cn$$

$$\frac{cn}{2} + \frac{cn}{2} = cn$$

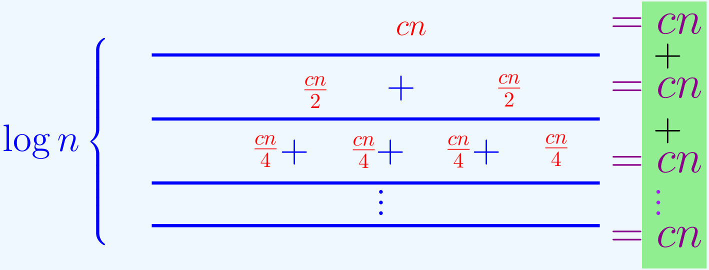
$$\frac{cn}{4} + \frac{cn}{4} + \frac{cn}{4} + \frac{cn}{4} = cn$$

$$\vdots \qquad \vdots$$

$$= cn$$

Recursion Trees

An illustrated example...



$$= cn \log n = O(n \log n)$$

Merge Sort Variant

Question: Merge Sort splits into 2 (roughly) equal sized arrays. Can we do better by splitting into more than 2 arrays? Say k arrays of size n/k each?

10.7
Quick Sort

Quick Sort

Quick Sort [Hoare]

1. Pick a pivot element from array
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself. Linear scan of array does it. Time is $O(n)$
3. Recursively sort the subarrays, and concatenate them.

Quick Sort

Quick Sort [Hoare]

1. Pick a pivot element from array
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself. Linear scan of array does it. Time is $O(n)$
3. Recursively sort the subarrays, and concatenate them.

Quick Sort

Quick Sort [Hoare]

1. Pick a pivot element from array
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself. Linear scan of array does it. Time is $O(n)$
3. Recursively sort the subarrays, and concatenate them.

Quick Sort

Quick Sort [Hoare]

1. Pick a pivot element from array
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself. Linear scan of array does it. Time is $O(n)$
3. Recursively sort the subarrays, and concatenate them.

Quick Sort: Example

1. array: 16, 12, 14, 20, 5, 3, 18, 19, 1
2. pivot: 16

Time Analysis

1. Let k be the rank of the chosen pivot. Then,
   $T(n) = T(k-1) + T(n-k) + O(n)$

Time Analysis

1. Let k be the rank of the chosen pivot. Then,
   $T(n) = T(k-1) + T(n-k) + O(n)$
2. If $k = \lceil n/2 \rceil$ then
   $T(n) = T(\lceil n/2 \rceil - 1) + T(\lfloor n/2 \rfloor) + O(n) \leq 2T(n/2) + O(n)$. Then,
   $T(n) = O(n \log n)$.

Time Analysis

1. Let k be the rank of the chosen pivot. Then,
   $T(n) = T(k-1) + T(n-k) + O(n)$
2. If $k = \lceil n/2 \rceil$ then
   $T(n) = T(\lceil n/2 \rceil - 1) + T(\lfloor n/2 \rfloor) + O(n) \leq 2T(n/2) + O(n)$. Then,
   $T(n) = O(n \log n)$.
   2.1 Median can be found in linear time.

Time Analysis

1. Let k be the rank of the chosen pivot. Then,
   $T(n) = T(k-1) + T(n-k) + O(n)$
2. If $k = \lceil n/2 \rceil$ then
   $T(n) = T(\lceil n/2 \rceil - 1) + T(\lfloor n/2 \rfloor) + O(n) \leq 2T(n/2) + O(n)$. Then,
   $T(n) = O(n \log n)$.
   2.1 Median can be found in linear time.
3. Typically, pivot is the first or last element of array. Then,

$$T(n) = \max_{1 \leq k \leq n} (T(k-1) + T(n-k) + O(n))$$

   In the worst case $T(n) = T(n-1) + O(n)$, which means $T(n) = O(n^2)$.
   Happens if array is already sorted and pivot is always first element.

10.8
Binary Search

Binary Search in Sorted Arrays

Input Sorted array A of n numbers and number x

Goal Is x in A?

```
BinarySearch(A[a..b], x):
        if (b − a < 0) return NO
        mid = A[⌊(a + b)/2⌋]
        if (x = mid) return YES
        if (x < mid)
            return BinarySearch(A[a..⌊(a + b)/2⌋ − 1], x)
        else
            return BinarySearch(A[⌊(a + b)/2⌋ + 1..b], x)
```

Analysis: $T(n) = T(\lfloor n/2 \rfloor) + O(1)$. $T(n) = O(\log n)$.

Observation: After k steps, size of array left is $n/2^k$

Binary Search in Sorted Arrays

Input Sorted array A of n numbers and number x

Goal Is x in A?

```
BinarySearch(A[a..b], x):
    if (b − a < 0) return NO
    mid = A[⌊(a + b)/2⌋]
    if (x = mid) return YES
    if (x < mid)
        return BinarySearch(A[a..⌊(a + b)/2⌋ − 1], x)
    else
        return BinarySearch(A[⌊(a + b)/2⌋ + 1..b],x)
```

Analysis: $T(n) = T(\lfloor n/2 \rfloor) + O(1)$. $T(n) = O(\log n)$.

Observation: After k steps, size of array left is $n/2^k$

Binary Search in Sorted Arrays

Input Sorted array A of n numbers and number x

Goal Is x in A?

```
BinarySearch(A[a..b], x):
        if (b − a < 0) return NO
        mid = A[⌊(a + b)/2⌋]
        if (x = mid) return YES
        if (x < mid)
            return BinarySearch(A[a..⌊(a + b)/2⌋ − 1], x)
        else
            return BinarySearch(A[⌊(a + b)/2⌋ + 1..b], x)
```

Analysis: $T(n) = T(\lfloor n/2 \rfloor) + O(1)$. $T(n) = O(\log n)$.

Observation: After k steps, size of array left is $n/2^k$

Another common use of binary search

1. Optimization version: find solution of best (say minimum) value
2. Decision version: is there a solution of value at most a given value v?

Reduce optimization to decision (may be easier to think about):

1. Given instance I compute upper bound $U(I)$ on best value
2. Compute lower bound $L(I)$ on best value
3. Do binary search on interval $[L(I), U(I)]$ using decision version as black box
4. $O(\log(U(I) - L(I)))$ calls to decision version if $U(I), L(I)$ are integers

Another common use of binary search

1. Optimization version: find solution of best (say minimum) value
2. Decision version: is there a solution of value at most a given value v?

Reduce optimization to decision (may be easier to think about):

1. Given instance I compute upper bound $U(I)$ on best value
2. Compute lower bound $L(I)$ on best value
3. Do binary search on interval $[L(I), U(I)]$ using decision version as black box
4. $O(\log(U(I) - L(I)))$ calls to decision version if $U(I), L(I)$ are integers

Example

1. Problem: shortest paths in a graph.
2. Decision version: given G with non-negative integer edge lengths, nodes s, t and bound B, is there an s-t path in G of length at most B?
3. Optimization version: find the length of a shortest path between s and t in G.

Question: given a black box algorithm for the decision version, can we obtain an algorithm for the optimization version?

Example continued

Question: given a black box algorithm for the decision version, can we obtain an algorithm for the optimization version?

1. Let U be maximum edge length in G.
2. Minimum edge length is L.
3. s-t shortest path length is at most $(n-1)U$ and at least L.
4. Apply binary search on the interval $[L,(n-1)U]$ via the algorithm for the decision problem.
5. $O(\log((n-1)U-L))$ calls to the decision problem algorithm sufficient. Polynomial in input size.

10.9
Solving Recurrences

Solving Recurrences

Two general methods:

1. Recursion tree method: need to do sums
   1.1 elementary methods, geometric series
   1.2 integration
2. Guess and Verify
   2.1 guessing involves intuition, experience and trial & error
   2.2 verification is via induction

Recurrence: Example I

1. Consider $T(n) = 2T(n/2) + n/\log n$ for $n > 2$, $T(2) = 1$.
2. Construct recursion tree, and observe pattern. ith level has 2^i nodes, and problem size at each node is $n/2^i$ and hence work at each node is $\frac{n}{2^i}/\log\frac{n}{2^i}$.
3. Summing over all levels

$$
\begin{aligned}
T(n) &= \sum_{i=0}^{\log n - 1} 2^i \left[\frac{(n/2^i)}{\log(n/2^i)}\right] \\
&= \sum_{i=0}^{\log n - 1} \frac{n}{\log n - i} \\
&= n \sum_{j=1}^{\log n} \frac{1}{j} = nH_{\log n} = \Theta(n \log\log n)
\end{aligned}
$$

Recurrence: Example I

1. Consider $T(n) = 2T(n/2) + n/\log n$ for $n > 2$, $T(2) = 1$.
2. Construct recursion tree, and observe pattern. ith level has 2^i nodes, and problem size at each node is $n/2^i$ and hence work at each node is $\frac{n}{2^i} / \log \frac{n}{2^i}$.
3. Summing over all levels

$$\begin{aligned}
T(n) &= \sum_{i=0}^{\log n - 1} 2^i \left[\frac{(n/2^i)}{\log(n/2^i)} \right] \\
&= \sum_{i=0}^{\log n - 1} \frac{n}{\log n - i} \\
&= n \sum_{j=1}^{\log n} \frac{1}{j} = nH_{\log n} = \Theta(n \log \log n)
\end{aligned}$$

Recurrence: Example II

1. Consider $T(n) = T(\sqrt{n}) + 1$ for $n > 2$, $T(2) = 1$.
2. What is the depth of recursion? $\sqrt{n}, \sqrt{\sqrt{n}}, \sqrt{\sqrt{\sqrt{n}}}, \ldots, O(1)$.
3. Number of levels: $n^{2^{-L}} = 2$ means $L = \log\log n$.
4. Number of children at each level is 1, work at each node is 1
5. Thus, $T(n) = \sum_{i=0}^{L} 1 = \Theta(L) = \Theta(\log\log n)$.

Recurrence: Example II

1. Consider $T(n) = T(\sqrt{n}) + 1$ for $n > 2$, $T(2) = 1$.
2. What is the depth of recursion? $\sqrt{n}, \sqrt{\sqrt{n}}, \sqrt{\sqrt{\sqrt{n}}}, \ldots, O(1)$.
3. Number of levels: $n^{2^{-L}} = 2$ means $L = \log\log n$.
4. Number of children at each level is 1, work at each node is 1
5. Thus, $T(n) = \sum_{i=0}^{L} 1 = \Theta(L) = \Theta(\log\log n)$.

Recurrence: Example III

1. Consider $T(n) = \sqrt{n}T(\sqrt{n}) + n$ for $n > 2$, $T(2) = 1$.
2. Using recursion trees: number of levels $L = \log\log n$
3. Work at each level? Root is n, next level is $\sqrt{n} \times \sqrt{n} = n$. Can check that each level is n.
4. Thus, $T(n) = \Theta(n \log\log n)$

Recurrence: Example III

1. Consider $T(n) = \sqrt{n}T(\sqrt{n}) + n$ for $n > 2$, $T(2) = 1$.
2. Using recursion trees: number of levels $L = \log\log n$
3. Work at each level? Root is n, next level is $\sqrt{n} \times \sqrt{n} = n$. Can check that each level is n.
4. Thus, $T(n) = \Theta(n \log\log n)$

Recurrence: Example IV

1. Consider $T(n) = T(n/4) + T(3n/4) + n$ for $n > 4$. $T(n) = 1$ for $1 \leq n \leq 4$.
2. Using recursion tree, we observe the tree has leaves at different levels (a lop-sided tree).
3. Total work in any level is at most n. Total work in any level without leaves is exactly n.
4. Highest leaf is at level $\log_4 n$ and lowest leaf is at level $\log_{4/3} n$
5. Thus, $n \log_4 n \leq T(n) \leq n \log_{4/3} n$, which means $T(n) = \Theta(n \log n)$

Recurrence: Example IV

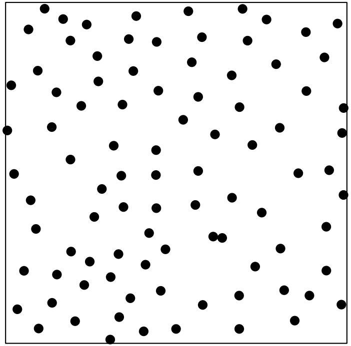
1. Consider $T(n) = T(n/4) + T(3n/4) + n$ for $n > 4$. $T(n) = 1$ for $1 \leq n \leq 4$.
2. Using recursion tree, we observe the tree has leaves at different levels (a lop-sided tree).
3. Total work in any level is at most n. Total work in any level without leaves is exactly n.
4. Highest leaf is at level $\log_4 n$ and lowest leaf is at level $\log_{4/3} n$
5. Thus, $n \log_4 n \leq T(n) \leq n \log_{4/3} n$, which means $T(n) = \Theta(n \log n)$

10.10
Supplemental: Divide and conquer for closest pair

Problem: Closest pair

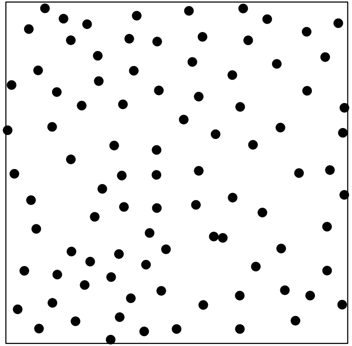
P: Set of n distinct points in the plane.
Compute the two points $p, q \in P$ that are closest together. Formally, compute

$$\arg \min_{p,q \in P : p \neq q} ||p - q||.$$



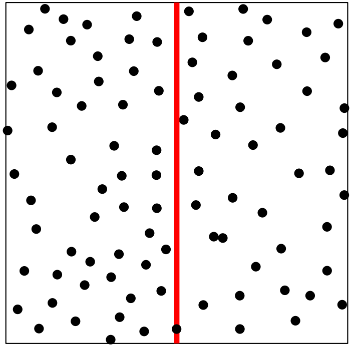
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$

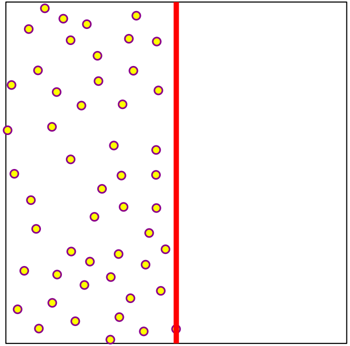


Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$



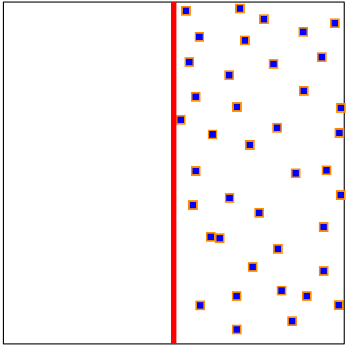
Closest pair: Divide and conquer leads to a special case



1. $P = P_L \cup P_R$

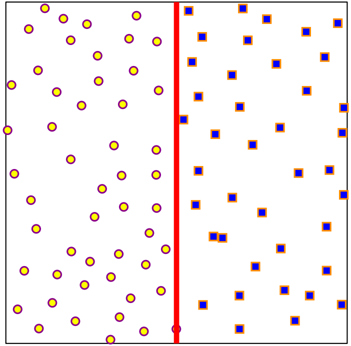
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$



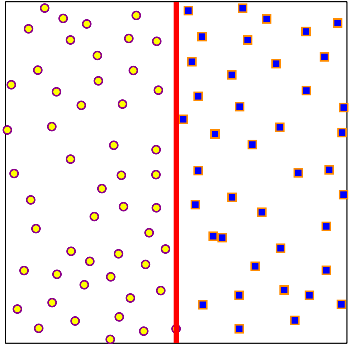
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$
2. $|P_L| = |P_R| = n/2$.
   $x(P_L) < 0$ and $x(P_R) > 0$.



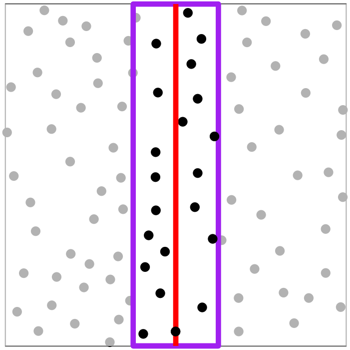
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$
2. $|P_L| = |P_R| = n/2$.
   $x(P_L) < 0$ and $x(P_R) > 0$.
3. Given $\ell = \min(\mathrm{cp}(P_L), \mathrm{cp}(P_R))$.



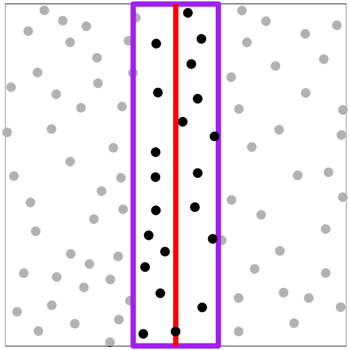
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$
2. $|P_L| = |P_R| = n/2$.
   $x(P_L) < 0$ and $x(P_R) > 0$.
3. Given $\ell = \min(\mathrm{cp}(P_L), \mathrm{cp}(P_R))$.
4. $P_m = \{p \in P \mid -\ell \le x(p) \le \ell\}$



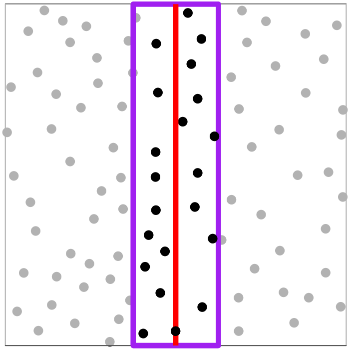
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$
2. $|P_L| = |P_R| = n/2$.
   $x(P_L) < 0$ and $x(P_R) > 0$.
3. Given $\ell = \min(\mathrm{cp}(P_L), \mathrm{cp}(P_R))$.
4. $P_m = \{p \in P \mid -\ell \leq x(p) \leq \ell\}$
5. Task: compute
   $\mathrm{cp}(P) = \min(\ell, \mathrm{cp}(p_M))$.



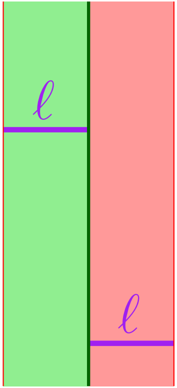
Closest pair: Divide and conquer leads to a special case

1. $P = P_L \cup P_R$
2. $|P_L| = |P_R| = n/2$.
   $x(P_L) < 0$ and $x(P_R) > 0$.
3. Given $\ell = \min(\mathrm{cp}(P_L), \mathrm{cp}(P_R))$.
4. $P_m = \{p \in P \mid -\ell \le x(p) \le \ell\}$
5. Task: compute
   $\mathrm{cp}(P) = \min(\ell, \mathrm{cp}(p_M))$.
6. **Claim:** Closest pair in P_m can be computed in $O(n \log n)$ time.



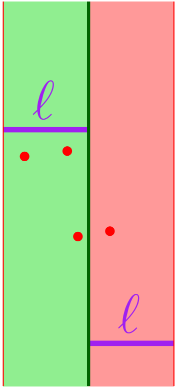
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



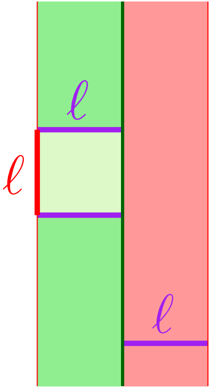
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



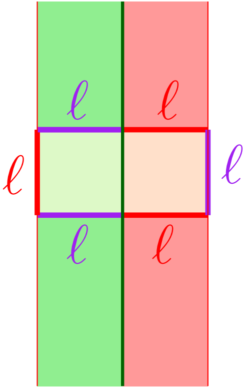
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



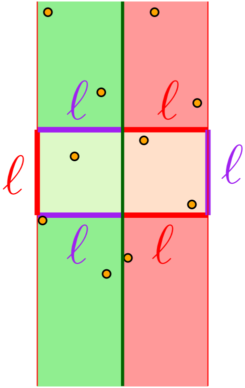
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



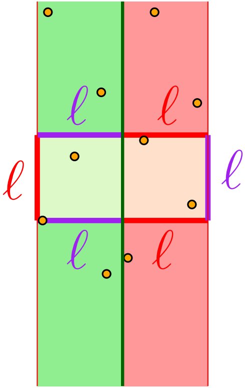
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



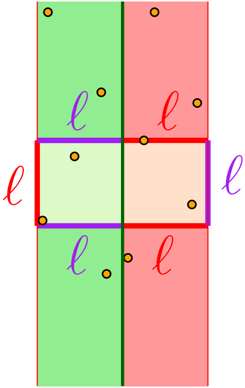
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



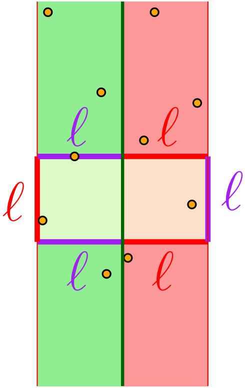
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



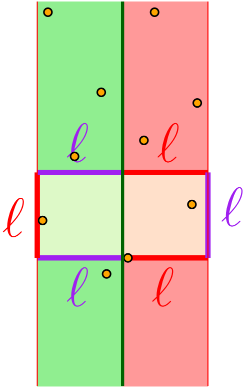
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



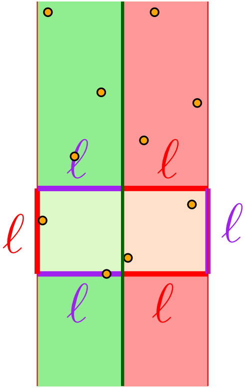
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



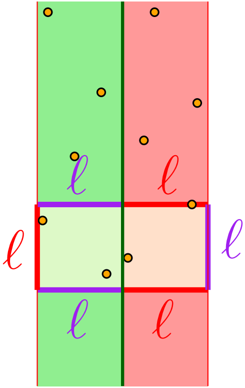
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



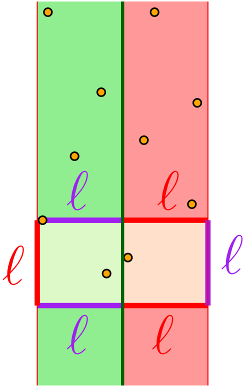
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



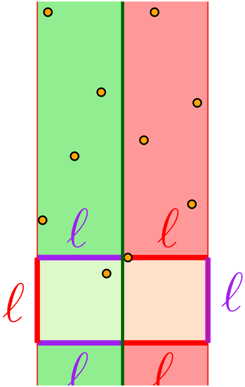
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



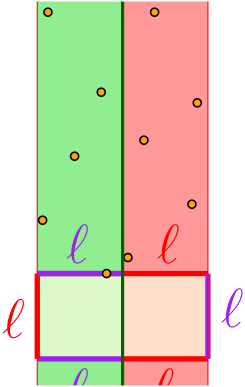
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



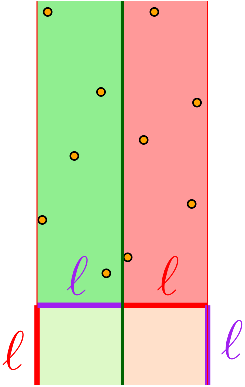
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



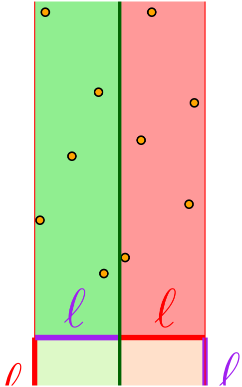
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



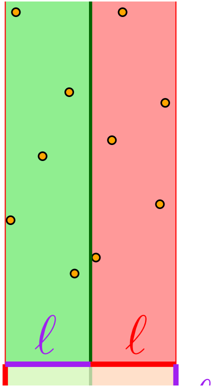
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



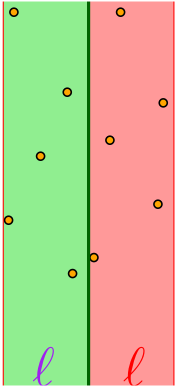
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



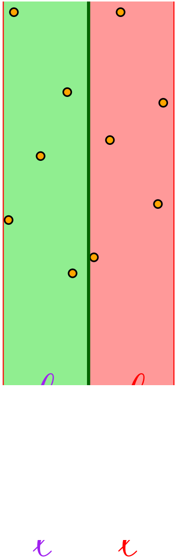
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



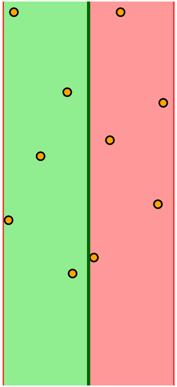
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



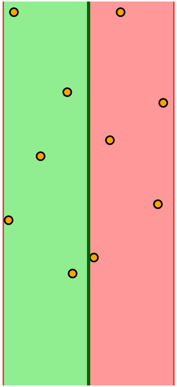
An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



An elevator can not be too full

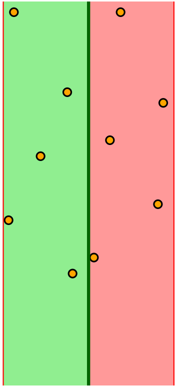
...or $\mathbf{P_m}$ is well spread



Closet pair in P_m can be computed in $O(n \log n)$ time.

An elevator can not be too full

...or $\mathbf{P_m}$ is well spread



Closet pair in P_m can be computed in $O(n \log n)$ time.

Closet pair in P_m can be computed in $O(n)$ time, if P is presorted by y-order.

Implementing the elevator trip upward

- Input:
  - Input: Points $p_1, p_2 \ldots, p_m$ sorted in inc order by y.
  - ℓ: Current minimum closest-pair distance encountered.
  - $\forall i \qquad x(p_i) \in [\text{median} - \ell, \text{median} + \ell]$.
- Elevator: Range $[\![b : t]\!] = \{b, b+1, \ldots, t\}$
- Points in elevator in any point in time: $p_b, p_{b+1}, \ldots, p_t$.
- $t, b \leftarrow 1$
- **while** $t \leq m$ **do**

  ```
  // If bottom point in elevator too far to be candidate
  // to be closest pair with top point, then throw it away.
  ```

  while $y(p_b) < y(p_t) - \ell$ **do** $b++$

  for $i \in [\![b : t-1]\!]$ **do** $\ell = \min(\ell, ||p_t - p_i||)$.
- **return** ℓ

Closest pair: Algorithm

CPDInner = **ClosestPairDistance**

```
CPDInner( P = {p_1, ..., p_n} ):
    if |P| = O(1) then compute by brute force
    x* = median(x(p_1), ..., x(p_n)).
    P_L ← {p ∈ P | x(p) ≤ x*}
    P_R ← {p ∈ P | x(p) > x*}
    ℓ_L = CPDInner(P_L)
    ℓ_R = CPDInner(P_R)
    ℓ = min(ℓ_L, ℓ_R).
    P_m = {p ∈ P | x* − ℓ ≤ x(p) ≤ x* + ℓ}
    ℓ_M = call alg.  closet-pair distance for special case on P_m.
    return min(ℓ, ℓ_m).
```

```
CPD( P = {p_1, ..., p_n} ):
    return CPDInner(P)
```

Closest pair algorithm

Lemma 10.1.

Given a set P of n points in the plane, one can compute the closet pair distance in P in $O(n \log^2 n)$ time.

Closest pair: Algorithm

CPDInner = **ClosestPairDistance**

```
CPDInner( P = {p_1, ..., p_n} ):
    if |P| = O(1) then compute by brute force
    x* = median(x(p_1), ..., x(p_n)).
    P_L ← {p ∈ P | x(p) ≤ x*}
    P_R ← {p ∈ P | x(p) > x*}
    ℓ_L = CPDInner(P_L)
    ℓ_R = CPDInner(P_R)
    ℓ = min(ℓ_L, ℓ_R).
    P_m = {p ∈ P | x* − ℓ ≤ x(p) ≤ x* + ℓ}
    ℓ_M = call alg. closet-pair distance for special case on P_m.
    return min(ℓ, ℓ_m).
```

```
CPD( P = {p_1, ..., p_n} ):
    Sort P by x-order. Sort P by y-order
    return CPDInner(P)
```

Closest pair

Theorem 10.2.

Given a set P of n points in the plane, one can compute the closet pair distance in P in $O(n \log n)$ time.

Wait wait... one can do better

Rabin showed that if we allow the floor function, and randomization, one can do better:

Theorem 10.3.

Given a set P of n points in the plane, one can compute the closet pair distance in P in $O(n)$ time.