Part I

Brief Intro to Algorithm Design and Analysis

Algorithms and Computing

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

A computer is a device that

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

Models of Computation vs Computers

- Model of Computation: an idealized mathematical construct that describes the primitive instructions and other details.
- 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 the Turing Machine model.
Unit-Cost RAM Model

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

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.
- For some problems such as basic arithmetic computation, unit-cost model makes no sense. Examples: multiplication of two $n$-digit numbers, primality etc.
- 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.
- 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.

Example

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

We will usually not allow (or be careful about allowing):
- bitwise operations (and, or, xor, shift, etc).
- floor function.
- limit word size (usually assume unbounded word size).

Models used in class

In this course when we design algorithms:
- Assume unit-cost RAM by default.
- We will explicitly point out where unit-cost RAM is not applicable for the problem at hand.
- 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.
**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 $A$ solves $f$ if for all input strings $w$, $A$ 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.

**Analysis of Algorithms**

Given a problem $P$ and an algorithm $A$ for $P$ we want to know:

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

**Asymptotic running-time analysis:** $A$ runs in $O(f(n))$ time if:

“For all $n$ and for all inputs $I$ of size $n$, $A$ on input $I$ terminates after $O(f(n))$ primitive steps.”

**Types of Problems**

We will broadly see three types of problems.

- **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 $w \in L(G)$?

- **Search Problem:** Find a solution if input is a YES input.
  - Example: Given graph $G$, nodes $s$, $t$, find an $s$-$t$ path.

- **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.

**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
Part II
Reductions and Recursion

Reduction
Reducing problem $A$ to problem $B$:
- 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 he turns blue, and then shoot him 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:

```plaintext
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

```plaintext
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

- **Positive direction**: Algorithm for B implies an algorithm for A
- **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

- An $O(n \log n)$ time algorithm for Sorting implies an $O(n \log n)$ time algorithm for Distinct Elements problem.
- 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.

Maximum Independent Set in a Graph

**Definition**

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 $(u, v) \not\in E$.

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.

- 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!

```plaintext
MaxWeightIndependentArcs(arcs C)
    cur-max = 0
    for each arc C \in 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
```
Recursion

- Recursion is a very powerful and fundamental technique
- Basis for several other methods
  - Divide and conquer
  - Dynamic programming
  - Enumeration and branch and bound etc
  - Some classes of greedy algorithms
- Makes proof of correctness easy (via induction)
- Recurrences arise in analysis

Tower of Hanoi

Move stack of $n$ disks from peg 0 to peg 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

Recursive Algorithm

\[
Hanoi(n, \text{src}, \text{dest}, \text{tmp}):\nonumber
\begin{align*}
\text{if } (n > 0) & \text{ then} \\
Hanoi(n-1, \text{src}, \text{tmp}, \text{dest}) & \quad \text{Move disk } n \text{ from } \text{src} \text{ to } \text{dest} \\
Hanoi(n-1, \text{tmp}, \text{dest}, \text{src}) & \quad \text{end recursion}
\end{align*}
\]

\[T(n): \text{ time to move } n \text{ disks via recursive strategy} \]

\[
T(n) = 2T(n-1) + 1 \quad n > 1 \quad \text{and } T(1) = 1
\]

Analysis

\[
T(n) = 2T(n-1) + 1 \\
= 2^2 T(n - 2) + 2 + 1 \\
= \ldots \\
= 2^i T(n - i) + 2^{i-1} + 2^{i-2} + \ldots + 1 \\
= \ldots \\
= 2^{n-1} T(1) + 2^{n-2} + \ldots + 1 \\
= 2^{n-1} + 2^{n-2} + \ldots + 1 \\
= (2^n - 1)/(2 - 1) = 2^n - 1
\]

Part III

Divide and Conquer
Divide and Conquer Paradigm

Divide and Conquer is a common and useful type of recursion.

**Approach**
- Break problem instance into smaller instances - divide step
- Recursively solve problem on smaller instances
- Combine solutions to smaller instances to obtain a solution to the original instance - conquer step

**Question:** Why is this not plain recursion?
- 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.
- There are many examples of this particular type of recursion that it deserves its own treatment.

Merge Sort [von Neumann]

- **Input:** Array $A[1 \ldots n]$

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

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

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

- Recursively MergeSort $A[1 \ldots m]$ and $A[m + 1 \ldots n]$

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

- Merge the sorted arrays

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

Sorting

**Input** Given an array of $n$ elements

**Goal** Rearrange them in ascending order

Merging Sorted Arrays

- Use a new array $C$ to store the merged array

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

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

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

\[
\text{MergeSort}(A[1..n]):
\begin{align*}
&\text{if } n > 1 \\
&\quad \text{m } \leftarrow \lfloor n/2 \rfloor \\
&\quad \text{MergeSort}(A[1..m]) \\
&\quad \text{MergeSort}(A[m+1..n]) \\
&\quad \text{Merge}(A[1..n], m)
\end{align*}
\]

| Merge(A[1..n], m): | \\
|---|---|
| i \leftarrow 1; j \leftarrow m + 1 | \\
| for k \leftarrow 1 to n |
| if j > n |
| \quad B[k] \leftarrow A[i]; i \leftarrow i + 1 |
| else if i > m |
| \quad B[k] \leftarrow A[j]; j \leftarrow j + 1 |
| else |
| \quad B[k] \leftarrow A[i]; i \leftarrow i + 1 |
| else |
| \quad B[k] \leftarrow A[j]; j \leftarrow j + 1 |
| for k \leftarrow 1 to n |
| A[k] \leftarrow B[k] |

---

Running Time

\[ T(n): \text{time for merge sort to sort an } n \text{ element array} \]

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

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

- \( T(n) = O(f(n)) \) - upper bound
- \( T(n) = \Omega(f(n)) \) - lower bound

---

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.

At the start of iteration \( k \) the following hold:

- \( B[1..k] \) contains the smallest \( k \) elements of \( A \) correctly sorted.
- \( B[1..k] \) contains the elements of \( A[1..(i - 1)] \) and \( A[(m + 1)..(j - 1)] \).
- No element of \( A \) is modified.

---

Solving Recurrences: Some Techniques

- Know some basic math: geometric series, logarithms, exponentials, elementary calculus
- Expand the recurrence and spot a pattern and use simple math
- Recursion tree method — imagine the computation as a tree
- 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

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?

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

- array: 16, 12, 14, 20, 5, 3, 18, 19, 1
- pivot: 16
Let $k$ be the rank of the chosen pivot. Then,

$$T(n) = T(k - 1) + T(n - k) + O(n)$$

If $k = \lceil n/2 \rceil$ then

$$T(n) = T(\lceil n/2 \rceil - 1) + T(\lceil n/2 \rceil) + O(n) \leq 2T(n/2) + O(n).$$

Then, $T(n) = O(n \log n)$.

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.

**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
- $mid = A[\lceil (a + b)/2 \rceil]$:
  - if ($x = mid$) return YES
  - if ($x < mid$)
    - return BinarySearch($A[a..\lceil (a + b)/2 \rceil - 1]$, $x$)
  - else
    - return BinarySearch($A[\lceil (a + b)/2 \rceil + 1..b]$, $x$)

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

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

**Another common use of binary search**

- **Optimization version:** find solution of best (say minimum) value
- **Decision version:** is there a solution of value at most a given value $v$?

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

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

- Problem: shortest paths in a graph.
- 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 $B$?
- 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

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

Part V

Solving Recurrences

Two general methods:
- Recursion tree method: need to do sums
  - elementary methods, geometric series
  - integration
- Guess and Verify
  - guessing involves intuition, experience and trial & error
  - verification is via induction
Recurrence: Example I

- Consider $T(n) = 2T(n/2) + n/\log n$ for $n > 2$, $T(2) = 1$.
- Construct recursion tree, and observe pattern. $i$th level has $2^i$ nodes, and problem size at each node is $n/2^i$ and hence work at each node is $n/2^i \log n$.
- Summing over all levels
  $$T(n) = \sum_{i=0}^{\log n - 1} 2^i \left[ \frac{(n/2^i)}{\log(n/2^i)} \right]$$
  $$= \sum_{i=0}^{\log n - 1} \frac{n}{\log n - i}$$
  $$= n \sum_{j=1}^{\log n} \frac{1}{j} = n H_{\log n} = \Theta(n \log \log n)$$

Recurrence: Example II

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

Recurrence: Example III

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

Recurrence: Example IV

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