International Institute of
Information Technology, Hyderabad

Algorithmic Theory

Data Structures, Algorithms and Competitive Programming

Algorithms Notes

Maintainers: Animesh Sinha, Gaurang Tandon, Dixit Garg, Arpan Dasgupta.

June 25, 2021

Chapter 1

Problem Solving Strategies

1.1 General Advice and Basic Patters

- If you need to Minimize a Maximum or Maximize a Minimum, use Binary Search.

  1. ICPCkh19-ANAJOBS

- If you want to use DP / Expectation to find something optimal, use Exchange Argument.

  1. ICPCgw17-KALADIN

1.2 Standard Problems for Practice

1.2.1 Math and Geometry

- Substitutions can help reduce the dimentionality of the problem (Kernel trick).

  1. CF1142-C: Make the Substitution $z_i = y_i^2 - x_i$. Now the problem is just computing the upper convex hull.

1.3 Trees and Graphs

- Substitutions can help reduce the dimentionality of the problem (Kernel trick).

  1. CF1142-C: Nice use of LCA for finding.

- On shortest paths with pay-for-ticket/earn-by-staying, expand state space and use Dijkstra.

1. ABC164-E Simple idea with one constraint.
2. CF1341-E: Uses 0-1 BFS to speed up.

Chapter 2

Probability

2.1 Methods of Solving Problems

All Probability or Expectation Value problems are solvable using these techniques, so trying to classify all problems in one of the following is reasonable.

Probability times Value (The Definition)

Expectation by Definition is Probability times the Value of the event. If the state space is ennumerable, then this is the technique to go with. Sometimes, **Dynamic Programming** can be used with this technique.

Technique of Contributions (Exchange Argument)

For each element, find what is the Probability that it contributes to the total answer, and how much does it contribute. Try to solve for every pair of elements, then take it to all the elements.

Indicator Variables

Interconverting sums using indicator variables is often a hard and algebraically tedious topic, hence we mark it as a #TODO.

2.2 Problems and Solutions

2.2.1 Codeforces 100371-H: Granite of Science

Problem 2.1: Granite of Science

There are n subjects, each having n_i lessons. Let the lessons be indexed by j. We are looking at the first m days of the semester. Any subject can start on any of the m days (even if all it's lesson's don't fit till the end), and each lesson will follow consecutively and contigously. Given that $h_{i,j}$ is the fatigue due to i-th subject's j-th lesson, and effective fatigue in one day is the sum of squares of fatigues due to all lessons that day. Total Fatigue is sum of effective fatigue in each of m days. Compute expected total fatigue.

Tags : 2D Prefix Sum , Linearity of Expectation, Independent Events.

Following is one solution to this problem (simplest and cleanest). Another one using Indicator Variables also exists, but seems to overcomplicate a relatively simple question.

Notation: n is number of subjects. f_{day} is the sum of difficulties of all subjects on that day. $f_{\text{day}} = \sum_{i=1}^{i} d_i, j$, where d_i is difficulty of i-th subject for that particular day and j-th lesson.

$$
\begin{aligned}
\text{ans} &= \sum_{\text{day}=1}^{n} f_{\text{day}}^2 \\
\implies E[\text{ans}] &= E\left[\sum_{\text{day}=1}^{2m} f_{\text{day}}^2\right] \\
&= \sum_{\text{day}=1}^{2m} E\left[f_{\text{day}}^2\right]
\end{aligned}
$$

Now, $f_{\text{day}}^2 = (c_1 + c_2 + \ldots + c_n)^2 = (\sum_i c_i^2 + 2\sum_{i,j;i\neq j} c_i c_j)$

$$
\begin{aligned}
E\left[f_{\text{day}}^2\right] &= \sum_i E[c_i^2] + 2\sum_{i,j;i\neq j} E[c_i c_j] \\
&= \sum_i E[c_i^2] + 2\sum_{i,j;i\neq j} E[c_i]E[c_j]
\end{aligned}
$$

Because i-th and j-th subject's contribution on the same day are independent of each other, their expectations can be split. The final expression can be simplified using the following:

$$
2\sum_{i,j;i\neq j} E[c_i]E[c_j] = \sum_{i,j} E[c_i]E[c_j] - \sum_i E[c_i]^2
$$

Thus, now we can use prefix sums and solve this problem in $\mathcal{O}(N^2)$.

Chapter 3

Heaps

3.1 Question Patterns

3.1.1 Generic Types

- Find the kth Minimum or Maximum. This can also be on arrays or trees with online insertion or deletion.
- Priority Queue questions. Typically greedy algorithms best implemented this way. (eg. Dijkstra)

3.1.2 Specific Illustrations

When not to use Heaps *Question: Find the kth-Maximum sum of all subarrays of any given array.* This question demostrates that when we have **k-sorted elements appended to our queue of k-maximum elements** at one go, it's better to use an actual **Sorted array and the Merge Function for Updates**.

Lazy Deletion Heaps cannot search, so heaps cannot delete. A simple solution for this is to maintain another heap of all deleted elements, and if the actual heap and deleted heap have the same top element, keep popping them out. We only care about the top element, so we can be Lazy in the deletion of the elements lower down in the tree.

3.2 Elementary Theory

3.2.1 Basic Operations

Heapification A **O(n)** algorithm exists for Heapification. It uses the standard sift-down procedure, but starts by correcting the lowest layers first. So start at the end of the array and move back.

Insert A **O(log n)** algorithm is used to insert. We push the element at the end of the array and sift-up or sift-down the element till the element is at a valid locale.

Delete For a **O(log n)** deletion algorithm, we must swap the last element with the element to be deleted, sift-up/down the former last element, and pop-out the last element. This is only possible if we have a refernce to the element, search and delete take O(n) time to perform on heaps.

Minimum/Maximum Minimum on a MinHeap takes **O(1)** time, and so does Maximum on a MaxHeap.

Chapter 4

Range Queries

4.1 Square Root Decomposition

Break the array down into chunks of $\sqrt{n}$. Store the answer to these chunks. The answer to range queries is the answer in the starting chunk after L, plus in the ending chunk before R, plus the stored results of everything in between. Updates can be done in $O(\sqrt{n})$ steps by just updating the result of the chunk.

4.1.1 Motivating Examples

This cannot be solved by a Segment Tree (easily).

Example 1.1

Question: Given and array, support update and query operation for number of elements less than k in the given range.
Solution: Maintain sorted vector of each block of the square-root decomposition. Search over all blocks and find lowerbound(k), the sum of these counts gives the answer in $O(\sqrt{n}log(n))$. For update, just sort again, using insertion sort you get $O(\sqrt{n})$.

4.2 Mo's Algorithm

This is an algorithm to handle offline queries by sorting them and trying to cleverly avoid recomputing portions that were already solved for in the previous queries. (Direcly usable with associative, commutative, invertible operations).

Algorithm

Break down the array in blocks of size $\sqrt{n}$. Sort queries by starting block, them by ending position. The right pointer keeps moving forward for each block, the left pointer keeps moving back and forth withing a block, adding elements as it goes back and subtracting as it goes forth. This results in all queries being solved in $O(n\sqrt{n})$ time.

Example 2.1

Question: Given an array, find the number of elements distinct elements in range (l, r). Sort the queries first by the starting block, then by the end position. Maintian a frequency array of all elements currently between right and left pointers. For each of the $O(\sqrt{n})$ blocks, the start pointer moves atmost $O(\sqrt{n})$ times back and forth in the block, adding and deleting elements. For each block the right pointer only goes forward adding in elements, $O(\sqrt{n})$ blocks each taking $O(n)$ time.

Example 2.2

Question: Given an array, find the f(s)*f(s)*s for all distinct s in range (l, r), where f(s) is the frequency of s.
Same Algorithm, and same frequency array as above, just find $f(s)*f(s)*s$ instead of $\delta(f(s))$ as above.

Chapter 5

Dynamic Programming

The focus of this chapter would be to enlist all the optimizations to a DP possible and types of recurrences solvable using them.

5.1 Matrix Exponentiation

Example 1.0: DP by Matrix Exponentiation

Number of ways to construct an array starting in X, and ending in Y, with no two adjacent elements are the same.

$$dp[i] = \begin{bmatrix} dp[i][\text{CLASH}] \\ dp[i][\text{CLEAN}] \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ k-1 & k-2 \end{bmatrix} \times dp[i-1]$$

5.2 Bitmasks

5.2.1 Classical Bitmasks

The idea is trivial, we take each bitmask to represent an arbitrary subset of any given set.

5.2.2 Sum over Subsets

We want to sum for each mask, some given function (as an array) for all it's submasks. We can take each mask to start off, and then go down a series of all it's subsets in decreasing order of value of mask, the algorithm for this will be $j = (j-1)\ \&\ mask$, initially $j_0 = mask$.

Time complexity is the following:

$$T(n) = \sum_{k=0}^{n-1} C_k^n 2^k = 3^k \tag{5.1}$$

Chapter 6

Game Theory

6.1 NIM Games and Sprague-Grundy Theorem

6.2 Take Away Games

6.2.1 Identifying the Losing States

Theorem 2.1

Let H_i denote all the losing states, and $f(x)$ denote the number of stones that can be removed in the next move after x stones in the previous. Then we can find the losing states as follows.

$$H_{k+1} = H_k + H_m, \quad where \;\; m = min\{j : f(H_j) \geq H_k\} \tag{6.1}$$

The idea is that we can remove any $H_j + H_k$ stones, we can think of them as two separate piles. We cannot win on either pile, so the only way to win is when the H_j pile ends, the last move was enough that $f(last\,move) \geq H_k$ so that we can win next move. If this is not possible, then the state is losing.

KEY IDEA: Find the RECURRENCE, make a SOLUTION HYPOTHESIS by monitoring the pattern and prove it BY INDUCTION to get all the losing states.

6.2.2 A few Example Functions

Example: f(x) = x

1 stone is losing, so $H_1 = 1$. And whenever H_k is losing, the $min\{H_j : f(H_j) \geq H_k\} = min\{H_j : H_j \geq H_k\} = H_k$, therefore the losing states are 2^n.

Example: f(x) = 2x - 1

1 stone is losing, so $H_1 = 1$. Our Hypothesis, $H_k = 2H_{k-1}$. And whenever H_k is losing, the $min\{H_j : f(H_j) \geq H_k\} = min\{H_j : 2 * H_j - 1 \geq H_k\} = H_k$, therefore the losing states are 2^n.

Example: f(x) = 2x: Fibonacci NIM

1 stone is losing, so $H_1 = 1$. Our Hypothesis, $H_k = H_{k-1} + H_{k-2}$. And whenever H_k is losing, the $min\{H_j : f(H_j) \geq H_k\} = min\{H_j : 2 * H_j \geq H_k\} = H_{k-1}$, therefore the losing states are the Fibonacci numbers.

6.2.3 Winning Strategy

New Binary number systems

We find that we can express any number as a sum of the values of H_1, H_2, ..., so we construct a binary like number system where a the place value of the i-th digit is H_i and the face value is 0 or 1. Let's call this H-binary. (Note: This expression is unique and complete for powers of 2, and for the Fibonacci numbers - Zeckendorf theorem, as in the above examples).

The greedy strategy

Given any starting state that is not losing, we can write out it's representation in the H-Binary system. Since this will have more than 1 ones in it's representation, we subtract the LEAST SIGNIFICANT BIT .

Now, the opponent cannot remove the next one in the representation, because of the property of number systems that $H_j > f(H_i) \;\; \forall j > i$, due to the way we found losing states H_i. Finally, when our oppnent removes any value from the form .....1000000 (Any value, last set bit 1, and 0s), he will get a 1 in the resulting representation .....0000110.

Now in our move we shall remove the lowest set bit again. This is possible, as the last move must have been greater than or equal to H_j if j is lowest set bit. (Obviously, because when we add back we need to have 1+1 = 0 to get all the numbers back to 0 and 1 at the position that could not be removed). So $f(H_j) >= H_j$, this move is possible, and we can win.

6.2.4 Proof of Victory

On our moves, we reduce the number of ones in the representation by 1. Our opponent, if he removes the one at H_j, he has to insert a 1 and position smaller than j. So he increases or keeps constant the number of ones. Obviously, the last move will be played by us, reducing the number of ones to 0 and finishing the game.

6.2.5 References

Problems

Fibonacci Nim (Direct Implementation) [ICPC Kolkata 2018] `https://www.codechef.com/KOL18ROL/problems/SNOWMAN`

Theory

Contains most of the theory mentioned above: `http://www.cut-the-knot.org/Curriculum/Games/TakeAway.shtml#theory`

6.3 Finding Invarients

Mark out a state and all it's children. Either try MINIMAX TREE, and if the state space is large, try to find invarients, specially MODULO or PARITY.

Example 3.0: Invarients in Games - 1

Question: Start with {(4 sticks, length 4), (1 stick, length 1)}. In a move, we can break a stick or remove k sticks of length k. Last move wins. Find the winning states & strategy. Any state be (n1, n2, n3, n4). All states with (n1 + n3) % 2 == (n2 + n4) % 3 are winning positions, all others are losing. We can prove than any winning position goes to losing position and vice-versa.

Chapter 7

Mathematical Tools

7.1 Fast Fourier Transforms

7.1.1 Motivation and Purpose

We want to be able to interconvert a polynomial between **SAMPLES representation and COEFFICIENTS representation**.

A n-degree polynomial can be a n-dimentional COEFFICIENTS vector where it is easy to compute the value at any random x, and as a n-dimentional SAMPLES vector that has n (x, y) pairs, making it easy to multiply vectors. If we can convert back and forth in $O(n \cdot log(n))$, and *multiply polynomials* in SAMPLES land in $O(n)$, then we get a speed up over the typical $O(n^2)$ for multipling each coefficient with every other.

7.1.2 Algorithm

The Recurrence

Given a polynomial $P(x)$, to convert it into samples at $X = \{X_1, X_2 ..., X_n\}$. We can solve it using the recurrence

$$P(X) = P_{even}(x^2) + x \cdot P_{odd}(x^2) \tag{7.1}$$

where P_{even} is the polynomials with only even coefficients, P_{odd} with only odd. Note that the degree of the resulting polynomials P_{even} and P_{odd} is half of the original.

The Complex Numbers

We also want the size of the set X, at which we have to evaluate the polynomial to go down. So we can use complex numbers, for polynomial of degree n, we use **the 2^k-roots of unity** where 2^k is the smallest value power of 2 bigger than n.

The set will keep collapsing to half it's size after each step, as squares of exactly two of these roots is the same.

The Divide and Conquer

In summary, we are performing a Divide and Conquer solve, where each state is ¡P, X¿. We start with our **Original P, and X = 1**. In each divide step, we split the P into two parts, and each part gets one root of X, here +1 and -1. Then that divides in 4, at values +i, -i, +1 and -1. There are log(n) layers with 2^l polynomials of size $n/2^l$, each polynomial is to be evaluated at 1 value.

7.1.3 Some Mathematical Representations

$$\begin{bmatrix} 1 & x_1^2 & x_1^3 & x_1^4 & ... & x_1^n \\ 1 & x_2^2 & x_2^3 & x_2^4 & ... & x_2^n \\ 1 & x_3^2 & x_3^3 & x_3^4 & ... & x_3^n \\ ... & ... & ... & ... & ... & ... \\ 1 & x_n^2 & x_n^3 & x_n^4 & ... & x_n^n \end{bmatrix} \times \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ ... \\ c_n \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ ... \\ s_n \end{bmatrix} \tag{7.2}$$

Here we have the Vandermond Matrix of a set of N values for X times the coefficient vector gives the samples vector. We decided to choose our values in X as such: $X = \{1, \omega_n, \omega_n^2, \omega_n^3, ..., \omega_n^{n-1}\}$.

7.1.4 Inverse Fourier Transform

$$\begin{bmatrix} 1 & 1 & 1 & 1 & ... & 1 \\ 1 & \omega_n^1 & \omega_n^2 & \omega_n^3 & ... & \omega_n^{n-1} \\ 1 & \omega_n^2 & \omega_n^4 & \omega_n^6 & ... & \omega_n^{2n-2} \\ ... & ... & ... & ... & ... & ... \\ 1 & \omega_n^{n-1} & \omega_n^{2n-2} & \omega_n^{3n-3} & ... & \omega_n^{(n-1)(n-1)} \end{bmatrix}^{-1} \times \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ ... \\ s_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ ... \\ c_n \end{bmatrix} \tag{7.3}$$

Here is the definition of the operation. We are going from samples to coefficient, so we need multiplication by inverse of the matrix for Fourier. This is easy, because the inverse is just the complex conjugate divided by n.

$$V^{-1} = \bar{V}/n \tag{7.4}$$

So we can use the FFT Algorithm again, since $X = 1, \bar{\omega}^1, \bar{\omega}^2, \bar{\omega}^3, ..., \bar{\omega}^n$. And divide the answer by n. Note that X is still the same set, so no change to FFT is needed.

7.1.5 Number Theoretic Transforms

The number theoretic transform is based on generalizing the n-th primitive root of unity to a "quotient ring" instead of the usual field of complex numbers.

We take a number w that satisfies $w^n \equiv 1 (modp)$ going through each of the numbers only and atmost once.

7.2 Group Theory

7.2.1 Burnside's Lemma

It states the number of elements in the Orbit of a When a Group G acts on a Set X is the mean of the number of unique elements in the subgroup due to

$$|X/G| = \frac{1}{|G|}\sum_{g \in G} |X^g| \tag{7.5}$$

where X^g is the number of elements in set X fixed by the element g,

7.3 Number Theory

7.3.1 Stern-Brocot Tree

It is a **Binary Search Tree** of Fractions such than the path from the root to any number, is an incrementally closer set of approximations (Continued Fraction approximations) to that number.

Notation: we represent the continued fraction as an array.

$$a_0 + \frac{1}{a_1 + \frac{1}{\ldots + \frac{1}{k}}} = [a_0; a_1, ..., a_k]$$

This representation is not unique, since $[a_0; a_1, ..., a_k] = [a_0, a_1, ..., a_k - 1, 1]$ because $\frac{1}{a_k} = \frac{1}{(a_k - 1) + \frac{1}{1}}$

Parent-Child Relations

Parent of $[a_0, a_1, ..., a_k]$ is $[a_0, a_1, ..., a_k - 1]$
Children of $[a_0, a_1, ..., a_k]$ are $[a_0, a_1, ..., a_k - 1, 2]$ and $[a_0, a_1, ..., a_k + 1]$.

7.3.2 Chinese Remainder Theorem

System $x \equiv a_i (\mod m_i)$ for $i = 1, \ldots, n$, with pairwise relatively prime m_i has a unique solution modulo $M = \prod m_i$ $x = \sum_i a_i b_i \frac{M}{m_i} (\mod M)$ where b_i is modular inverse of $\frac{M}{m_i}$ modulo m_i.

System $x \equiv a (\mod m)$, $x \equiv b (\mod n)$ has solutions iff $a \equiv b (\mod g)$, where $g = \gcd(m, n)$. The solution is unique modulo $L = \frac{mn}{g}$, and equals: $x \equiv a + T(b-a)m/g \equiv b + S(a-b)n/g (\mod L)$, where S and T are integer solutions of $mT + nS = \gcd(m, n)$.

Theorems

Euler's theorem: $a^{\phi(n)} \equiv 1 (\mod n)$, if $\gcd(a, n) = 1$ Wilson's theorem: p is prime iff $(p-1)! \equiv -1 (\mod p)$ Primitive Pythagorean triple generator: $(m^2 - n^2)^2 + (2mn)^2 = (m^2 + n^2)^2$ Postage stamps/McNuggets problem: Let

a, b be coprime integers. There are exactly $\frac{1}{2}(a-1)(b-1)$ numbers not of form $ax+by(x,y \geq 0)$, and the largest is $(a-1)(b-1)-1 = ab-a-b$.

Fermat's two-squares theorem: Odd prime p can be represented as a sum of two squares iff $p \equiv 1(\mod 4)$. A product of two sums of two squares is a sum of two squares. Thus, n is a sum of two squares iff every prime of form $p = 4k+3$ occurs an even number of times in n's factorization.

Counting Primes Fast: To count number of primes lesser than big n. Use following recurrence. $\text{dp}[n][j] = \text{dp}[n][j+1] + \text{dp}[n/p_j][j]$ where $dp[i][j]$ stores count of numbers lesser than equal to i having all prime divisors greater than equal to p_j . Precompute this for all i less than some small k and for others use the recurrence to compute in small time.

Compute $P_N(x)$ in:

$$T(n) = T(n/2) + \mathcal{O}(n \log n) \tag{7.6}$$

$P_{2N}(x) = P_N(x)P_N(x+N)$, using polynomial shifting. Say, $P_N(x) = \prod\limits_{i=1}^{N}(x+i) = \sum_{i=0}^{N} c_i.x^i$. Then, $P_N(x+N) = \sum_{i=0}^{N} h_i.x^i$, where, $h_i = \frac{1}{i!}.(\text{coefficient of } x^{N-i} in A(x)B(x))$ where, $A(x) = \sum\limits_{i=0}^{N}(c_{N-i}.(N-i)!).x^i$, and $B(x) = \sum\limits_{i=0}^{N}\left(\frac{N^i}{i!}\right).x^i$

```
MUL(N) // computes (x+1)(x+2)...(x+N) in O(NlogN)
    if N==1:
        return (x+1)
    C = MUL(N/2)
    H = convolute(A,B) // use C to obtain A
    ANS = convolute(C,H)
    if N is odd:
        ANS *= (x+N) // naive multiplication will do - O(N)
    return ANS
```

Computing 10^{18}-th Fib number fast: use $f(2k) = f(k)^2 + f(k-1)^2, f(2k+1) = f(k)f(k+1) + f(k-1)f(k)$. This has at most $\mathcal{O}(\log n \log\log n)$ states.

7.3.3 Mobius Inversions

- $\phi \circ I = \text{id}$ i.e. $\sum_{d|n} \phi(d) = n$. Hence, $\phi = \mu \circ \text{id}$ i.e. $\phi(d) = \sum_{d|n} \mu(d)\frac{n}{d}$
- Count of numbers coprime to n and lesser than $n = phi(n)$
  Sum of numbers coprime to n and lesser than n is $\frac{n}{2}\phi(n)$
  Proved using the fact that if x is coprime to n then so is $n-x$ coprime to n. Sum over both and take average
- $\sum_{d|n} \mu(d)f(d) = \prod_{p|n}(1-f(p))$ (p are its prime factors)
- $\sum_{d|n} \mu^2(d)f(d) = \prod_{p|n}(1+f(p))$
- $\phi(mn) = \frac{\phi(m)\phi(n)\gcd(m,n)}{\phi(\gcd(m,n))}$

- $\phi(p^a) = p^{a-1}\phi(p)$

Chapter 8

Graphs and Tree

8.1 Trees

8.1.1 Heavy Light Decomposition

Motivating Problem

Given a Tree, handle the following queries:

- Update(edge, weight): Change the weight of any given edge in the Tree.
- Query(nodeX, nodeY): Find the heaviest eadge between two nodes, x and y.

8.2 Basic Algorithms on Graphs

8.2.1 List of algorithms

- Depth First Search
- Breadth First Search
- Shortest Path - Dijkstra's
- Shortest Path - Bellman Ford
- Shortest Path - Floyd Warshall
- Connected Components
- Topological Sort
- Prim's Maximum Spanning Tree

Chapter 9

Flow Algorithms

9.1 Some interesting types of Flows

9.1.1 Circulation Flows

When there is no source, but we need to find if there exists a valid cyclic flow in a graph with directed edges with minimum and maximum capacities. The flow through each pipe has to be more than minimum and less than maximum.

9.2 Chains in Posets

A Poset (Partially Ordered Set) is a set where $<$, $=$, $>$ are defined between some pair of elements but not over others. A subset where comparators are defined between every pair is called a chain, where none is defined is called an anti-chain.

Theorem 2.1: Dilworth's Theorem

The maximum length of any anti-chain in a poset equals the number of chains in it's minimal chain cover.

Theorem 2.2: Koenig's Theorem

The minimum set cover of a bipartite graph equals the maximum flow from one set to the other.