Seoul National University

Swift Turtwig

Joonik Bae, Minkyu Son, Wonseok Shin

2021-07-29

Data structures (1)

1.1 GCC Extension Data Structures

OrderStatisticTree.h

HashMap.h

1.2 Interval Tree structures

SegmentTree.h

LazySegmentTree.h

2DSegmentTree.h

FenwickTree.h

FenwickTree2d.h

MergeSortTree.h

1.3 Li-Chao Tree

LiChaoTree.h

1.4 Miscellaneous

UnionFind.h

UnionFindRollback.h

Treap.h

SubMatrix.h

MoQueries.h

Mathematics (2)

2.1 Geometry

2.1.1 Triangles

Side lengths: a, b, c

Semiperimeter: $p = \dfrac{a+b+c}{2}$

Area: $A = \sqrt{p(p-a)(p-b)(p-c)}$

Circumradius: $R = \dfrac{abc}{4A}$

Inradius: $r = \dfrac{A}{p}$

Length of median (divides triangle into two equal-area triangles): $m_a = \frac{1}{2}\sqrt{2b^2 + 2c^2 - a^2}$

Length of bisector (divides angles in two):

$$s_a = \sqrt{bc\left[1 - \left(\frac{a}{b+c}\right)^2\right]}$$

Law of sines: $\dfrac{\sin\alpha}{a} = \dfrac{\sin\beta}{b} = \dfrac{\sin\gamma}{c} = \dfrac{1}{2R}$

Law of cosines: $a^2 = b^2 + c^2 - 2bc\cos\alpha$

Law of tangents: $\dfrac{a+b}{a-b} = \dfrac{\tan\frac{\alpha+\beta}{2}}{\tan\frac{\alpha-\beta}{2}}$

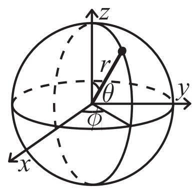
2.1.2 Quadrilaterals

With side lengths a, b, c, d, diagonals e, f, diagonals angle θ, area A and magic flux $F = b^2 + d^2 - a^2 - c^2$:

$$4A = 2ef \cdot \sin\theta = F\tan\theta = \sqrt{4e^2f^2 - F^2}$$

For cyclic quadrilaterals the sum of opposite angles is 180°, $ef = ac + bd$, and $A = \sqrt{(p-a)(p-b)(p-c)(p-d)}$.

2.1.3 Spherical coordinates



$$x = r\sin\theta\cos\phi \qquad r = \sqrt{x^2+y^2+z^2}$$
$$y = r\sin\theta\sin\phi \qquad \theta = \operatorname{acos}(z/\sqrt{x^2+y^2+z^2})$$
$$z = r\cos\theta \qquad \phi = \operatorname{atan2}(y, x)$$

2.2 Matrices

Matrix.h

Determinant.h

SolveLinear.h

SolveLinearBinary.h

MatrixInverse.h

2.3 FFT, Berlekamp

FastFourierTransform.h

NumberTheoreticTransform.h

BerlekampMassey.h

Number Theory (3)

3.1 Primes

```
< 10^k    prime   # of prime
----------------------------
1             7            4
2            97           25
3           997          168
4          9973         1229
5         99991         9592
6        999983        78498
7       9999991       664579
8      99999989      5761455
9     999999937     50847534
```

Primitive roots exist modulo any prime power p^a, except for $p = 2, a > 2$, and there are $\phi(\phi(p^a))$ many. For $p = 2, a > 2$, the group $\mathbb{Z}_{2^a}^\times$ is instead isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_{2^{a-2}}$.

3.2 Estimates

$\sum_{d|n} d = O(n \log \log n)$.

The number of divisors of n is at most around 100 for $n < 5e4$, 500 for $n < 1e7$, 2000 for $n < 1e10$, 200 000 for $n < 1e19$.

3.3 Modular arithmetic

ModInverse.h

ModPow.h

ModLog.h

ModSum.h

ModSqrt.h

3.4 Primality

FastEratosthenes.h

MillerRabin.h

Factor.h

3.5 Divisibility

euclid.h

CRT.h

3.5.1 Bézout's identity

For $a \neq, b \neq 0$, then $d = gcd(a, b)$ is the smallest positive integer for which there are integer solutions to

$$ax + by = d$$

If (x, y) is one solution, then all solutions are given by

$$\left(x + \frac{kb}{\gcd(a,b)}, y - \frac{ka}{\gcd(a,b)}\right), \quad k \in \mathbb{Z}$$

3.6 Fractions

ContinuedFractions.h

FracBinarySearch.h

3.7 Mobius Function

$$\mu(n) = \begin{cases} 0 & n \text{ is not square free} \\ 1 & n \text{ has even number of prime factors} \\ -1 & n \text{ has odd number of prime factors} \end{cases}$$

Mobius Inversion:

$$g(n) = \sum_{d|n} f(d) \Leftrightarrow f(n) = \sum_{d|n} \mu(d) g(n/d)$$

Other useful formulas/forms:

$\sum_{d|n} \mu(d) = [n = 1]$ (very useful)

$g(n) = \sum_{n|d} f(d) \Leftrightarrow f(n) = \sum_{n|d} \mu(d/n) g(d)$

$g(n) = \sum_{1 \le m \le n} f(\lfloor \frac{n}{m} \rfloor) \Leftrightarrow f(n) = \sum_{1 \le m \le n} \mu(m) g(\lfloor \frac{n}{m} \rfloor)$

Numerical (4)

4.1 Polynomials and recurrences

Polynomial.h

PolyRoots.h

PolyInterpolate.h

4.2 Optimization

IntegrateAdaptive.h

Simplex.h

Combinatorial (5)

5.1 Permutations

5.1.1 Factorial

n	1 2 3	4	5	6	7	8	9	10
$n!$	1 2 6	24	120	720	5040	40320	362880	3628800

n	11	12	13	14	15	16	17
$n!$	4.0e7	4.8e8	6.2e9	8.7e10	1.3e12	2.1e13	3.6e14

n	20	25	30	40	50	100	150	171
$n!$	2e18	2e25	3e32	8e47	3e64	9e157	6e262	>DBL_MAX

IntPerm.h

5.1.2 Cycles

Let $g_S(n)$ be the number of n-permutations whose cycle lengths all belong to the set S. Then

$$\sum_{n=0}^{\infty} g_S(n) \frac{x^n}{n!} = \exp\left(\sum_{n \in S} \frac{x^n}{n}\right)$$

5.1.3 Derangements

Permutations of a set such that none of the elements appear in their original position.

$$D(n) = (n-1)(D(n-1) + D(n-2)) = nD(n-1) + (-1)^n = \left\lfloor \frac{n!}{e} \right\rceil$$

5.1.4 Burnside's lemma

Given a group G of symmetries and a set X, the number of elements of X *up to symmetry* equals

$$\frac{1}{|G|} \sum_{g \in G} |X^g|,$$

where X^g are the elements fixed by g ($g.x = x$).

If $f(n)$ counts "configurations" (of some sort) of length n, we can ignore rotational symmetry using $G = \mathbb{Z}_n$ to get

$$g(n) = \frac{1}{n} \sum_{k=0}^{n-1} f(\gcd(n, k)) = \frac{1}{n} \sum_{k|n} f(k) \phi(n/k).$$

5.2 Partitions and subsets

5.2.1 Partition function

Number of ways of writing n as a sum of positive integers, disregarding the order of the summands.

$$p(0) = 1, \; p(n) = \sum_{k \in \mathbb{Z} \setminus \{0\}} (-1)^{k+1} p(n - k(3k-1)/2)$$

$$p(n) \sim 0.145/n \cdot \exp(2.56\sqrt{n})$$

n	0	1	2	3	4	5	6	7	8	9	20	50	100
$p(n)$	1	1	2	3	5	7	11	15	22	30	627	~2e5	~2e8

5.2.2 Lucas' Theorem

Let n, m be non-negative integers and p a prime. Write $n = n_k p^k + ... + n_1 p + n_0$ and $m = m_k p^k + ... + m_1 p + m_0$. Then $\binom{n}{m} \equiv \prod_{i=0}^{k} \binom{n_i}{m_i} \pmod{p}$.

5.2.3 Binomials

multinomial.h

5.3 General purpose numbers

5.3.1 Bernoulli numbers

EGF of Bernoulli numbers is $B(t) = \frac{t}{e^t-1}$ (FFT-able).
$B[0,\ldots] = [1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, \ldots]$

Sums of powers:

$$\sum_{i=1}^{n} n^m = \frac{1}{m+1}\sum_{k=0}^{m}\binom{m+1}{k}B_k \cdot (n+1)^{m+1-k}$$

Euler-Maclaurin formula for infinite sums:

$$\sum_{i=m}^{\infty} f(i) = \int_m^\infty f(x)dx - \sum_{k=1}^{\infty}\frac{B_k}{k!}f^{(k-1)}(m)$$

$$\approx \int_m^\infty f(x)dx + \frac{f(m)}{2} - \frac{f'(m)}{12} + \frac{f'''(m)}{720} + O(f^{(5)}(m))$$

5.3.2 Stirling numbers of the first kind

Number of permutations on n items with k cycles.

$$c(n,k) = c(n-1,k-1) + (n-1)c(n-1,k),\ c(0,0) = 1$$
$$\textstyle\sum_{k=0}^{n} c(n,k)x^k = x(x+1)\ldots(x+n-1)$$

$c(8,k) = 8, 0, 5040, 13068, 13132, 6769, 1960, 322, 28, 1$
$c(n,2) = 0, 0, 1, 3, 11, 50, 274, 1764, 13068, 109584, \ldots$

5.3.3 Eulerian numbers

Number of permutations $\pi \in S_n$ in which exactly k elements are greater than the previous element. k j:s s.t. $\pi(j) > \pi(j+1)$, $k+1$ j:s s.t. $\pi(j) \geq j$, k j:s s.t. $\pi(j) > j$.

$$E(n,k) = (n-k)E(n-1,k-1) + (k+1)E(n-1,k)$$

$$E(n,0) = E(n,n-1) = 1$$

$$E(n,k) = \sum_{j=0}^{k}(-1)^j\binom{n+1}{j}(k+1-j)^n$$

5.3.4 Stirling numbers of the second kind

Partitions of n distinct elements into exactly k groups.

$$S(n,k) = S(n-1,k-1) + kS(n-1,k)$$

$$S(n,1) = S(n,n) = 1$$

$$S(n,k) = \frac{1}{k!}\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}j^n$$

5.3.5 Bell numbers

Total number of partitions of n distinct elements. $B(n) = 1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, \ldots$. For p prime,

$$B(p^m + n) \equiv mB(n) + B(n+1) \pmod{p}$$

5.3.6 Labeled unrooted trees

\# on n vertices: n^{n-2}
\# on k existing trees of size n_i: $n_1 n_2 \cdots n_k n^{k-2}$
\# with degrees d_i: $(n-2)!/((d_1-1)!\cdots(d_n-1)!)$

5.3.7 Catalan numbers

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n+1} = \frac{(2n)!}{(n+1)!n!}$$

$$C_0 = 1,\ C_{n+1} = \frac{2(2n+1)}{n+2}C_n,\ C_{n+1} = \sum C_i C_{n-i}$$

$C_n = 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, \ldots$

- sub-diagonal monotone paths in an $n \times n$ grid.
- strings with n pairs of parenthesis, correctly nested.
- binary trees with with $n+1$ leaves (0 or 2 children).
- ordered trees with $n+1$ vertices.
- ways a convex polygon with $n+2$ sides can be cut into triangles by connecting vertices with straight lines.
- permutations of $[n]$ with no 3-term increasing subseq.

Graph (6)

6.1 Trees

LCA.h

HLD.h

LinkCutTree.h

DirectedMST.h

6.2 DFS algorithms

SCC.h

BiconnectedComponents.h

2sat.h

6.3 Euler walk

EulerWalk.h

6.4 Network flow

Dinic.h

MinCostMaxFlow.h

EdmondsKarp.h

6.5 Matching

hopcroftKarp.h

DFSMatching.h

MinimumVertexCover.h

WeightedMatching.h

GeneralMatching.h

6.6 Heuristics

MaximalCliques.h

MaximumClique.h

MaximumIndependentSet.h

Geometry (7)

7.1 Geometric primitives

Point.h

lineDistance.h

SegmentDistance.h

SegmentIntersection.h

lineIntersection.h

sideOf.h

OnSegment.h

linearTransformation.h

Angle.h

7.2 Circles

CircleIntersection.h

CircleTangents.h

CirclePolygonIntersection.h

circumcircle.h

MinimumEnclosingCircle.h

7.3 Polygons

InsidePolygon.h

PolygonArea.h

PolygonCenter.h

PolygonCut.h

ConvexHull.h

HullDiameter.h

PointInsideHull.h

LineHullIntersection.h

Checkpoints (10)

10.1 Debugging

- $10^5 * 10^5 \Rightarrow$ OVERFLOW. 특히 for 문 안에서 `i * i < n` 할때 조심하기.
- If unsure with overflow, use `#define int long long` and stop caring.
- 행렬과 기하의 i, j 인덱스 조심. 헷갈리면 쓰면서 가기.
- Segment Tree, Trie, Fenwick 등 Struct 구현체 사용할 때는 항상 내부의 n 이 제대로 초기화되었는지 확인하기.
- Testcase가 여러 개인 문제는 항상 초기화 문제를 확인하기. 입력을 다 받지 않았는데 break나 return으로 끊어버리면 안됨.
- iterator 주의 : .end() 는 항상 맨 끝 원소보다 하나 더 뒤의 iterator. erase쓸 때는 iterator++ 관련된 문제들에 주의.
- std::sort must compare with Strict weak ordering (Codejam 2020 1A-A)
- Memory Limit : Local variable은 int 10만개 정도까지만 사용. Global Variable의 경우 128MB면 대략 int 2000만 개까지는 잘 들어간다. long long은 절반. stack, queue, map, set 같은 특이한 컨테이너는 100만개를 잡으면 메모리가 버겁지만 vector 100만개는 잡아도 된다.
- Array out of Bound : 배열의 길이는 충분한가? Vector resize를 했다면 그것도 충분할까? 배열의 -1번에 접근한 적은 없는게 확실할까?
- Binary Search : 제대로 짠 게 맞을까? 1 차이 날 때 / lo == hi 일 때 등등. Infinite loop 주의하기.
- Graph : 반례 유의하기. Connected라는 말이 없으면 Disconnected. Acyclic 하다는 말이 없으면 Cycle 넣기, 특히 $A \leftrightarrow B$ 그래프로 2개짜리 사이클 생각하기.
- Set과 map은 매우 느리다.

10.2 Thinking

- 모든 경우를 다 할 수 없나? 왜 안 되지? 시간 복잡도 잘 생각해 보기. 정해의 Target Complexity를 먼저 생각하고 주요 알고리즘들의 Complexity로 짜맞추기.
  예를들어, 쿼리가 30만개 들어온다면 한 쿼리를 적어도 $\log n$ 에 처리할 방법이 아무튼 있다는 뜻.
- 그 방법이 뭐지? xxxx한 일을 어떤 시간복잡도에 실행하는 적절한 자료구조가 있다면?
  - 필요한 게 정렬성이라면 힙이나 map을 쓸 수 있고
  - multiset / multimap도 사용할 수 있고.. 느리지만.
- 단조함수이며, 충분히 빠르게 검증가능한가 : Binary Search.
- 차원이 높은 문제 : 차원 내려서 생각하기. 3 → 2. 2 → 1. 2019 Codejam R1B-1 Manhattaen Crepe Cart
- 이 문제가 사실 그래프 관련 문제는 아닐까?
  - 만약 그렇다면, '간선' 과 '정점' 은 각각..?
  - 간선과 정점이 몇 개 정도 있는가?
- 이 문제에 Overlapping Subproblem이 보이나?
  → Dynamic Programming 을 적용.
- Directed Graph, 특히 Cycle에 관한 문제 : Topological Sorting? (ex : SNUPC 2019 kdh9949)
- 답의 상한이 Reasonable 하게 작은가?
- output이 특정 수열/OX 형태 : 작은 예제를 Exhasutive Search. 모르는 뭔가를 알기 위해서는 데이터가 필요하다.
- 그래프 문제에서, 어떤 "조건" 이 들어갔을 때 → 이 문제를 "정점을 늘림으로써" 단순한 그래프 문제로 바꿀 수 있나? (ex : SNUPC 2018 달빛 여우) 이를테면, 홀짝성에 따라 점을 2배로 늘림으로써?
- DP도 마찬가지. 어떤 조건을 단순화하기 위해 상태의 수를 사이사이에 집어넣을 수 있나? (ex : SNUPC 2018 실버런)
- DP State를 어떻게 나타낼 것인가? 첫 i개만을 이용한 답을 알면 $i+1$개째가 들어왔을 때 빠르게 처리할 수 있을까?
- 더 큰 table에서 시작해서 줄여가기. 특히 Memory가 모자라다면 Toggling으로 차원 하나 내릴 수 있는 경우도 상당히 많이 있다. 각 칸의 갱신 시간과 칸의 개수 찾기.
- Square root Decomposition : $O(n \log n)$ 이 생각나면 좋을 것 같지만 잘 생각나지 않고, 제한을 보니 $O(n\sqrt{n})$ 이면 될것도 같이 생겼을 때 생각해 보기. $O(\sqrt{n})$ 버킷 테크닉. Red Army 2020 : Queue
- 복잡도가 맞는데 왜인지 안 뚫리면 : 필요없는 long long을 사용하지 않았나? map이나 set iterator 들을 보면서 상수 커팅. 간단한 함수들을 inlining. 재귀를 반복문으로. 특히 Set과 Map은 끔찍하게 느리다.
- 마지막 생각 : 조금 추하지만 해싱이나 Random 또는 `bitset` 을 이용한 $n^2/64$ 같은걸로 뚫을 수 있나? 컴파일러를 믿고 10^8의 몇 배 정도까지는 내 봐도 될 수도. 의외로 Naive한 문제가 많다. Atcoder 158 Divisible Substring