

# XCPC 算法竞赛模板

Sergio Gao

2025 年 8 月 19 日

# 目 录

<b>第 1 章 图论</b>	<b>3</b>
1.1 Tarjan 相关图论算法	3
1.1.1 求强连通分量	3
1.2 网络流	5
1.2.1 Dinic 最大流	5
1.2.2 Dijkstra 费用流	9
<b>第 2 章 凸优化相关</b>	<b>13</b>
2.1 四边形不等式相关	13
2.1.1 摘要	13
2.1.2 四边形不等式优化 DP	14
2.1.3 预备节: WQS 二分	15
2.1.4 WQS 二分结合决策单调性	20
2.2 维护凸包相关数据结构	22
2.2.1 李超线段树	22
<b>第 3 章 计算几何</b>	<b>30</b>
3.1 随机化方法	30
3.1.1 最小公共圆	30
3.2 数值算法	31
3.2.1 自适应 Simpson 积分	31
<b>第 4 章 数学</b>	<b>34</b>
4.1 初等数论相关	34
4.1.1 找模素数 $P$ 的原根	34
4.1.2 杜教筛求 $\text{premu}$	35
4.2 多项式相关	36
4.2.1 NTT 以及半在线卷积	36
<b>第 5 章 字符串</b>	<b>42</b>
5.1 自动机	42
5.1.1 AC 自动机	42

# 第 1 章 图论

主要记录一些比较常用的高级算法。

## 1.1 Tarjan 相关图论算法

叫 Tarjan 的图论算法其实有很多。所以 Tarjan 实际上可以说是一类思想。它可以说是洞察了 dfs 树的性质，考虑了返祖边，跨越边（对于无向图，实际上并不可能存在）等。

### 1.1.1 求强连通分量

这是对于有向图，求出若干极大两两可互相到达的点集。

一些没有写上去但可能会用到的 trick:

对于强连通分量，存在容易算法判断是否存在非零环。（提示：如果存在两个权值不一样的环，那么一定存在非零环。）

```
#include<bits/stdc++.h>
using namespace std;
class TarjanSCC
{
private:
    int n;
    vector<vector<int>> > e;
    vector<int> dfn, low;
    int dfncnt;
    stack<int> st;
    vector<bool> instack;
    void dfs(int u)
    {
        if(dfn[u]) return;
        low[u] = dfn[u] = ++dfncnt;
        st.push(u);
        instack[u] = 1;
        for(auto v: e[u])
        {
            if(!dfn[v])
            {
                dfs(v);
                low[u] = min(low[u], low[v]);
            }
        }
    }
};
```

```

        }
        else if(instack[v])
        {
            low[u] = min(low[u], dfn[v]);
        }
    }
    if(dfn[u] == low[u])
    {
        ++sc;
        while(!st.empty() && st.top() != u)
        {
            scc[st.top()] = sc;

            instack[st.top()] = 0;
            st.pop();
        }
        assert(!st.empty() && st.top() == u);
        scc[st.top()] = sc;
        instack[st.top()] = 0;
        st.pop();
    }
    return;
}

int sc;
public:
vector<int> scc;
TarjanSCC(int _n)
{
    n = _n;
    e = vector<vector<int>>>(n);
    dfn = vector<int>(n);
    low = vector<int>(n);
    st = stack<int>();
    instack = vector<bool>(n);
    scc = vector<int>(n);
    dfncnt = sc = 0;
}

void addEdge(int u, int v)
{
    assert(u < n && v < n);
    e[u].push_back(v);
    return;
} // point id should strictly less than n

void get()
{
    for(int i=0; i<n; i++) dfs(i);
    return;
}

```

```

};
void solve()
{
    int n, m;
    cin >> n >> m;
    TarjanSCC _graph(n+1);
    for(int i=1; i<=m; i++)
    {
        int u, v;
        cin >> u >> v;
        _graph.addEdge(u, v);
    }
    _graph.get();
    return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0);
    cout.tie(0);
    int T = 1;
    // cin >> T;
    while(T--)
    {
        solve();
    }
    return 0;
}

```

## 1.2 网络流

网络流在 XCPC 中常考的两类：

第一类是最大流或者最小割。这类一般建完图用 Dinic 算法解决即可。

特别地，最经典的一类题型：等价于求二分图的最大匹配。而最大匹配用最大流来跑又快又好，并且还能顺便证明二分图最大匹配 = 最小点覆盖。

二分图中：最大匹配（最大流）= 最小点覆盖（最小割）= 点数 - 最大独立集

第二类是最小费用最大流。这类一般建完图用 Dij 费用流解决即可。大多数费用流算法都和其流  $f$  相关。构造时一般是那最大流作为可以控制的约束。特别地，最经典的一类题型：等价于求二分图的最大权匹配。

### 1.2.1 Dinic 最大流

一般的时间复杂度会估计到  $O(n^2m)$ ，但很多题目的建图有特殊性，跑不满。特别的，如果拿来处理二分图最大匹配，复杂度可以估到  $O(n^3)$ 。

```

#include <bits/stdc++.h>
using namespace std;

```

```

template<class T> class Dinic
{
private:
    struct _Edge
    {
        int to;
        T cap;
    };
    int n;
    vector<_Edge> pool;
    vector<vector<int>> > g;
    vector<int> cur, h;
    bool bfs(int s, int t)
    {
        for(auto &i : h) i = -1;
        h[s] = 0;
        queue<int> q;
        q.push(s);
        while(!q.empty())
        {
            int u = q.front(); q.pop();
            for(int i=0; i<(int)g[u].size(); i++)
            {
                auto _to = pool[g[u][i]].to;
                auto _cap = pool[g[u][i]].cap;
                if( _cap == 0 || h[_to] !=-1 ) continue;
                h[_to] = h[u] + 1;
                if(_to == t) return true;
                q.push(_to);
            }
        }

        return false;
    }
    T dfs(int u, int t, T f)
    {
        if(u==t) return f;
        auto r = f;
        for(auto &i = cur[u]; i<(int)g[u].size(); i++)
        {
            auto _to = pool[g[u][i]].to;
            auto _cap = pool[g[u][i]].cap;
            if(h[_to] != h[u] + 1) continue;
            auto a = dfs(_to, t, min(r, _cap));
            pool[g[u][i]].cap -= a;
            pool[g[u][i]^1].cap += a;
            r -= a;
            if(r==0) break;
        }
    }
};

```

```

    }
    return f-r;
}
public:
Dinic(int _n)
{
    n = _n;
    cur.resize(n);
    h.resize(n);
    g.resize(n);
}
void addEdge(int u, int v, T c = numeric_limits<T>::max())
{
    g[u].push_back(pool.size());
    pool.push_back({v, c});

    g[v].push_back(pool.size());
    pool.push_back({u, 0});

    return;
}
T flow(int s, int t)
{
    T ans = 0;
    while(bfs(s, t))
    {
        for(auto &i : cur) i = 0;
        ans += dfs(s, t, numeric_limits<T>::max());
    }
    return ans;
}
};
void solve()
{
    int n, m, k;
    cin>>n>>m>>k;
    vector<int> p(m), invid(m);
    vector<pair<int, int> > pp(m);
    for(int i=0; i<m; i++) cin>>p[i], pp[i] = {p[i], i};
    sort(pp.begin(), pp.end());
    sort(p.begin(), p.end());
    for(int i=0; i<m; i++) invid[pp[i].second] = i;
    vector<int> s[n];
    for(int i=0; i<n; i++)
    {
        s[i] = vector<int>(m);
        for(int j=0; j<m; j++)
        {

```

```

        cin>>s[i][invid[j]];
    }
}
Dinic<long long> d(n*m + 1 + 1);
for(int i=0; i<n; i++)
{
    d.addEdge(0, i*m+0+1, s[i][0]);
    for(int j=0; j<m-1; j++)
    {
        // i*m+j+1 [i,j]
        d.addEdge(i*m+j+1, i*m+ j+1 +1, s[i][j+1]);
    }
    d.addEdge(i*m+(m-1)+1, n*m+1);
}
for(int i=1; i<=k; i++)
{
    int x, y, dis;
    cin>>x>>y>>dis;
    x--;
    y--;
    if(x==y) continue;
    for(int j=0, r=0; j<m; j++)
    {
        while(r<m && p[r] - p[j] <= dis) r++;
        if(r==m) break;
        d.addEdge(x*m+ (r-1) +1, y*m+ j +1);
        d.addEdge(y*m+ (r-1) +1, x*m+ j +1);
    }
}
cout<<d.flow(0, n*m+1)<<"\n";
return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0), cout.tie(0);

    int T;
    cin>>T;
    while(T--)
    {
        solve();
    }
    return 0;
}

```



### 1.2.2 Dijkstra 费用流

首先，费用流一般是要自己构造的，所以一般有意义的问题不会有真正的负环。然而负权是允许的。通过调整正负，可以把最小费用流的算法改写成最大费用流。

Dijkstra 费用流也就是在保证初始能赋出一个合理的势能函数的前提下，用 Dijkstra 来找到从源点到汇点的最短路。

最暴力的做法是初始做一遍 spfa，这个复杂度可能会达到  $O(nm)$ 。或者根据题目的特殊性。有一些问题它给出时就已经是让我们求如何最大化一些势能差之和。

除去这个初始化，复杂度是  $O(fm \log m)$ 。要注意，像是二部图最大权匹配， $f$  取决于较小的那一部，所以不要误认为其等于总点数。

```
#include<bits/stdc++.h>
using namespace std;
template<class T> class mcmf
{
private:
    const T INF = numeric_limits<T>::max();
    struct _edge
    {
        int to, cap;
        T cost;
    };
    int n;
    vector<_edge> e;
    vector<vector<int>> > g;
    int src, sink;
    vector<T> h;
    vector<T> dis;
    vector<pair<int, int>> > prev;
    bool Dij()
    {
        dis = vector<T>(n, INF);
        dis[src] = 0;
        priority_queue<pair<T, int>> > q;
        q.push({-dis[src], src});
        while(!q.empty())
        {
            auto frt = q.top();
            q.pop();
            if(frt.first != -dis[frt.second]) continue;
            int u = frt.second;
            for(int i=0; i<(int)g[u].size(); i++)
            {
                auto edge = e[g[u][i]];
                if(edge.cap > 0 && dis[edge.to] > dis[u] + h[u] + edge.cost - h[edge.to])
                {
                    dis[edge.to] = dis[u] + h[u] + edge.cost - h[edge.to];
                    prev[edge.to] = {u, g[u][i]};
                }
            }
        }
    }
};
```

```

        q.push({-dis[edge.to], edge.to});
    }
}
}
return dis[sink] < INF;
}
void spfa()
{
    queue<int> q;
    vector<int> vis(n, 0);
    h[src] = 0;
    vis[src] = 1;
    q.push(src);
    while(!q.empty())
    {
        int u = q.front();
        q.pop();
        vis[u] = 0;
        for(int i=0; i<(int)g[u].size(); i++)
        {
            auto edge = e[g[u][i]];
            if( edge.cap > 0 && h[edge.to] > h[u] + edge.cost)
            {
                h[edge.to] = h[u] + edge.cost;
                if( !vis[edge.to] )
                {
                    vis[edge.to] = 1;
                    q.push(edge.to);
                }
            }
        }
    }
    return;
}
public:
mcmf(int _n, int _src, int _sink)
{
    n = _n;
    src = _src;
    sink = _sink;
    g = vector<vector<int>> >(n);
    return;
}
void addEdge(int u, int v, int f, T c)
{
    g[u].push_back( e.size() );
    e.push_back({v, f, c});
    g[v].push_back( e.size() );
}

```

```

        e.push_back({u, 0, -c});
        return;
    }
    pair<int, T> get()
    {
        prev = vector<pair<int, int>>(n);
        int maxf = 0;
        T minc = 0;
        h = vector<T>(n, 0);
        spfa();
        while( Dij() )
        {
            int flow = numeric_limits<int>::max();
            for(int i=0; i<n; i++) if(dis[i] < INF) h[i] += dis[i];
            for(int v = sink; v != src; v = prev[v].first )
            {
                flow = min(flow, e[prev[v].second].cap);
            }
            for(int v = sink; v != src; v = prev[v].first )
            {
                e[prev[v].second].cap -= flow;
                e[prev[v].second^1].cap += flow;
            }
            maxf += flow;
            minc += h[sink] * flow;
        }
        return {maxf, minc};
    }
};

void solve()
{
    int n, m, s, t;
    cin>>n>>m>>s>>t;
    s--; t--;
    auto _graph = mcmf<int>(n, s, t);
    while(m--)
    {
        int u, v, w, c;
        cin>>u>>v>>w>>c;
        u--;
        v--;
        _graph.addEdge(u, v, w, c);
    }
    auto ans = _graph.get();
    cout<<ans.first << "_" << ans.second<<"\n";
    return;
}

```

```
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0);
    cout.tie(0);
    int T = 1;
    // cin>>T;
    while(T--)
    {
        solve();
    }
    return 0;
}
```

## 第 2 章 凸优化相关

目前看起来在 ICPC 区域赛几乎是每场必出。

### 2.1 四边形不等式相关

#### 2.1.1 摘要

要讨论的是这样的一类问题

$$f(j) = \min_{i \leq j} w(i, j)$$

关于某些  $w$  满足四边形不等式的证明：

(i) 把不等式写出来后，由于实质上只是左端点交换了一下，只与左端点有关的或者只与右端点有关的并不会改变，两边长得一样的，可以消去；常量更加可以消去了。for example, 有时候  $w(i, j)$  可能会被替换成  $w'(i, j) = dp[i-1] + w(i, j)$ ，然而  $dp[i-1]$  这一项只与左端点有关，所以其是否满足四边形不等式并不会被改变。还有就像是  $\lambda(j-i+1)$  这种可以被拆成左、右端点独立作用的也会被消去。

(ii) 通过一点小技巧，可以发现  $i < i+1 \leq j < j+1$  for  $j = j_0, j_0+1$  满足四边形不等式  $\implies i < i+1 \leq j < j+2$  满足四边形不等式，同理可以如此扩展左端点。所以可以只证明左端点相差 1 与右端点相差 1 的情况。

(iii) 对于显式的、可求二阶导的下凸函数  $f$ ， $w(i, j) := f(j-i)$ ，证明则不需要这么麻烦。稍微推导可以发现无非是要说明对于  $a < b \leq c < d$  at the same time  $a+d = b+c$ ，which actually represent the length of the 4 mentioned intervals, we have  $f(b) + f(c) \leq f(a) + f(d)$ 。FYI, there's a special interpretation on "intersect  $\leq$  include", which is "equal  $\leq$  different"。这样的话其实就是要说明  $f(b) - f(a) \leq f(d) - f(c)$ ，然而  $b-a = d-c$ ，从而就是要说明  $f(a+\Delta) - f(a) \leq f(c+\Delta) - f(c)$ 。这几乎就是下凸性质的定义。

(iv) **决策单调性优化 dp 学习笔记 — Exber's Blog** 上述性质似乎可以扩展（但没有提供证明），如果  $w(l, r)$  满足区间包含单调性和四边形不等式， $f(x)$  为下凸函数，则：

- $f(w(l, r))$  满足四边形不等式
- 如果  $f(x)$  单调不降（也就是没有触底反弹那一段），则  $f(w(l, r))$  还满足区间包含单调性

(v) 还有一些问题，在考虑四边形不等式进行代数运算时，不妨考虑对于三个区间  $[l_1, l_2], [l_2, r_1], [r_1, r_2]$  分别设一些变量，这样代数证明会明显一点。比如说证明

$$f(s_1 + s_2) + f(s_2 + s_3) \leq f(s_1 + s_2 + s_3) + f(s_2)$$

对于区间划分型 dp，听说恰经过  $k$  条边蒙日矩阵最短路关于  $k$  是下凸的。

## 2.1.2 四边形不等式优化 DP

### 2157. 「POI2011 R1」避雷针 Lightning Conductor - 题目 - LibreOJ

这个情景是，下标决策点未必是要小于当前点的。这种问题只需要正向跑一边，逆向跑一边，在两者之间取更优即可。模板是解决前面标准的

- 四边形不等式指相交  $\leq$  包含
- 等价于要决策的是最小值

注意中途运算时  $\text{sqrt}$  相关操作不能取整，反之则会破坏  $-\sqrt{x}$  的下凸性（比如说相邻两顶点值一样斜率就直接为 0 了）。

```
#include<bits/stdc++.h>
using namespace std;
const int N = 1e6 + 5;
int n, a[N], ans[N];
int ceilsq[N];
double f(int i, int j)
{
    return a[j] - a[i] + -sqrt(j-i);
}
void work()
{
    deque<pair<int, pair<int, int> > > q;
    for(int i=1; i<=n; i++)
    {
        if(!q.empty() && q.front().second.second < i ) q.pop_front();
        if(!q.empty() ) q.front().second.first = i;
        while( !q.empty() &&
            f(i, q.back().second.first) <= f(q.back().first, q.back().second.first) )
            q.pop_back();
        if(q.empty()) q.push_back({i, {i, n}});
        else if( f(i, n) < f(q.back().first, n) )
        {
            int l = q.back().second.first, r = n, ans = n;
            while(l<=r)
            {
                int mid = (l+r)/2;
                if( f(i, mid) < f(q.back().first, mid) )
                {
                    ans = mid;
                    r = mid-1;
                }
                else l = mid+1;
            }
            q.back().second.second = ans-1;
            q.push_back({i, {ans, n}});
        }
        ans[i] = min(ans[i], (int)floor(f( q.front().first, i) ) );
    }
}
```

```

    }
    return;
}
void solve()
{
    cin>>n;
    for(int i=1; i<=n; i++) cin>>a[i], ans[i] = 0;
    work();
    reverse(a+1, a+n+1);
    reverse(ans+1, ans+n+1);
    work();
    reverse(ans+1, ans+n+1);
    reverse(a+1, a+n+1);
    for(int i=1; i<=n; i++) cout<< (-ans[i]) <<"\n";
    return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0);
    cout.tie(0);
    int now = 0;
    for(int i=0; i<N; i++)
    {
        ceilsq[i] = now;
        if( now*now == i) now++;
    }
    int T = 1;
    // cin>>T;
    while(T--)
    {
        solve();
    }
    return 0;
}

```

### 2.1.3 预备节：WQS 二分

#### 传统方法

#### P2619 [国家集训队]Tree I - 洛谷

洛谷题解质量有点差，很多都看不出和 wqs 二分有什么关系。

这样思考，设  $g_x$  是恰选中  $x$  条白边的最小权。那么  $f(x) := g_x - \lambda x$  就是给这  $x$  条白边的权值减去  $\lambda$  的惩罚代价后的权值，那么  $g_x - \lambda x$  的最小值在什么地方取到？这等价于提前给所有白边都减去  $\lambda$ ，然后跑 mst。这有什么意义呢？

考虑  $(need-1, g_{need-1})$  和  $(need, g_{need})$ ，当  $\lambda$  恰好等于这两个点的斜率时，如果能有下凸性，并且同取最小值的时候可以取  $x$  大的，这时候就会取到后面的那个作为最小值的  $x$  点（如果取到  $> need$  的其实同理

可以往回取)。

当  $\lambda$  大于这两个点的斜率时,  $< need$  的部分比大于等于  $\geq need$  的部分不优。

当  $\lambda$  大于这两个点的斜率时,  $< need$  的部分比大于等于  $\geq need$  的部分更优。

所以二分这个惩罚代价, 使得  $p = \max\{pos : f(pos) = \min_x\{f(x)\}\}$  位于  $\geq need$  部分, 那么必然通过调整使得  $f(need) = f(p)$ 。

```
#include<bits/stdc++.h>
using namespace std;
const int N = 1e5 + 5;
int fa[N];
int father(int x)
{
    if(fa[x] == x) return x;
    return fa[x] = father(fa[x]);
}
void merge(int x, int y)
{
    x = father(x);
    y = father(y);
    fa[x] = y;
}
vector< pair<pair<int, int>, pair<int, int> > > e;
int V, E, need, val;
bool check(int d)
{
    for(int i=1; i<=V; i++) fa[i] = i;
    vector< pair<pair<int, int>, pair<int, int> > > tmpe = e;
    for(int i=1; i<=E; i++) if(tmpe[i].first.second == 0) tmpe[i].first.first += d;
    sort(tmpe.begin()+1, tmpe.end());
    int cnt = 0;
    val = 0;
    for(int i=1; i<=E; i++)
    {
        int s = tmpe[i].second.first,
            t = tmpe[i].second.second,
            c = tmpe[i].first.first,
            col = tmpe[i].first.second;
        if(father(s) == father(t)) continue;
        if(col==0) cnt++;
        val += c;
        merge(s, t);
    }
    return cnt >= need;
}
void solve()
{
    cin>>V>>E>>need;
    e = vector< pair<pair<int, int>, pair<int, int> > >(E+1);
```



```

for(int i=1; i<=E; i++)
{
    cin>>e[i].second.first>>e[i].second.second>>e[i].first.first>>e[i].first.second;
    e[i].second.first++;
    e[i].second.second++;
}
int l = -100, r = 100, ans = 666;
while(l<=r)
{
    int mid = (l+r)/2;
    if(check(mid))
    {
        ans = mid;
        l = mid+1;
    }
    else r = mid-1;
}
check(ans);
cout<< val - need*ans <<"\n";
return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0);
    cout.tie(0);
    int T = 1;
    // cin>>T;
    while(T--)
    {
        solve();
    }
    return 0;
}

```

## 通用方法

下面以二维 wqs 二分的题目为例，介绍一种不需要知道最小值横坐标，也就不需要处理共线问题的写法。

### CF1799F

对于两维分别凸的函数，它的性质并不像一维的那么好。比如说对于一维情况，点  $x$  的斜率自然地表述成  $x$  和  $x-1$  对应的点所确定的直线的斜率；对于二维呢？是考虑  $(x, y)$  与  $(x-1, y)$  与  $(x, y-1)$  构成的切平面吗？然而这个切平面未必能确定其与  $(x-1, y-1)$  的关系。另外，在一维的时候我们遇到的共线问题，在高维会变得更复杂。总之这个升维并不是显然的。

这里需要引入一个不太容易想得到的对偶函数（Legendre 变换）：对于上凸的  $f(x)$ ，令

$$g(k) := \max_x \{f(x) - k(x - a)\}$$

则可证明  $g(k)$  是下凸函数，且最小值在  $k = f'(a)$  时（离散情况则取离散定义，可能是一个区间，比如说  $[f(x+1) - f(x), f(x) - f(x-1)]$  中每一个  $k$ ）取到，即  $f(a)$ 。

这个操作的直观是，对于一个 fixed 的  $k$ ，在上凸壳的不同的点去做这个斜率的直线，交到  $x = a$  的位置中，取最大的。实际上我们会发现，这个最大位置实际上取的就是直线与上凸壳的切点。

只有整个上凸壳的切线恰好在经过  $(a, f(a))$  时，这个最大值才会取到下界  $f(a)$ 。为什么关于  $k$  会下凸呢？因为随着切点远离  $x$ ， $dk$  乘上这个距离会越来越大。

上面这种想法可以扩展。对于两维联合（这里只是另一维为常数意义下的话应该不太对的）上凸的  $f(x, y)$ ，现在想要单点求值， $f(a, b)$ 。那么只需要考虑由  $k_1, k_2$  构成的切平面， $z = k_1x + k_2y$ （这里省略了一个常数  $C$ ，我们会假想  $C$  使得恰好与二维上凸壳相切的，但其实整体平移没有影响）

$$g(k_1, k_2) := \max_{x, y} \{f(x, y) - k_1(x - a) - k_2(y - b)\}$$

应该能有  $g(k_1, k_2)$  关于两维联合下凸，且最小值在  $(k_1, k_2) = (f'_x(a, b), f'_y(a, b))$  处取到。（或者离散情况下，可以是包含  $(a, b)$  的某些平面的并）

联合下凸应该能保证，先预先对 fixed 的  $x$  找完  $y$  这维随便取后的最值，仍能在  $x$  方向有凸性。

上面的内容的严谨性还待我进一步阅读专业书籍。总之网上存在一些并不那么合适的论述，不当弱化了一些条件。

此外实现的时候不要忘记凸性取反。

```
#include<bits/stdc++.h>
using namespace std;

const int N = 5e3 + 5;

int n, a[N], b, k_1, k_2;

long long dp(long long lbd1, long long lbd2)
{
    long long res = 0;
    for(int i=1; i<=n; i++)
    {
        long long nw = res + a[i];
        nw = min(nw, res + (a[i]+1)/2 - lbd1);
        nw = min(nw, res + max(a[i]-b, 0) - lbd2);
        nw = min(nw, res + min( (max(a[i]-b, 0)+1)/2, max((a[i]+1)/2-b, 0) ) - lbd1 - lbd2);
        res = nw;
    }
    return res + lbd1 * k_1 + lbd2 * k_2;
}

long long check(long long lbd1)
{
    long long l = -2e9, r = 0, ans = -1e18;
    while(l<=r)
```

```

{
    long long mid = (l+r)/2;
    long long res = dp(lbd1, mid), _res = dp(lbd1, mid+1);
    ans = max(ans, max(res, _res));
    if(res <= _res) l = mid+2;
    else r = mid-1;
}
return ans;
}

void solve()
{
    cin>>n>>b>>k_1>>k_2;
    for(int i=1; i<=n; i++) cin>>a[i];

    long long l = -2e9, r = 0, ans = -1e18;
    while(l<=r)
    {
        long long mid = (l+r)/2;
        long long res = check(mid), _res = check(mid+1);
        ans = max(ans, max(res, _res));
        if(res <= _res) l = mid+2;
        else r = mid-1;
    }
    cout<<ans<<"\n";

    return;
}
/*

*/

int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0); cout.tie(0);

    int T = 1;
    cin>>T;
    while(T--) solve();
    return 0;
}
/*

*/

```

### 2.1.4 WQS 二分结合决策单调性

这个属于是固定套路，区间分划问题。下面给出对应模板代码情景的题解。

2024 Chengdu Regional F

考虑到

$$\left(\sum_{i=1}^m \frac{num_i}{k_i}\right) \left(\sum_{i=1}^m k_i sum_i\right) \geq \left(\sum_{i=1}^m \sqrt{num_i \cdot sum_i}\right)^2$$

因为要最小化  $\sum_{i=1}^m \frac{num_i}{k_i}$  所以用调整法可以说明  $(\sum_{i=1}^m k_i sum_i)$  会取其上界 1，这就把  $\{k\}$  给安排了。只需要最小化右式子。再用调整法可以说明 sum 里面应该是连续的，于是可以先对  $s$  排序，然后按顺序做一个 dp。这实际上就转化成了一个限制区间个数的区间分拆问题

WQS 二分，引入一个惩罚参数 lambda，将问题转化为：不限制分组数，但每多一组就要付出 lambda 的代价。可以看出惩罚参数越大分组越少，所以我们大概是要求一个适当大的惩罚参数使得分组的组数不超过  $m$ 。

这里的四边形不等式貌似不能直接套结论把  $\sqrt{x}$  剥掉直接考虑  $x$ ，因为这里就是要考虑  $w(i, j) = dp[i-1] + \sqrt{(j-i)(pre_j - pre_i)} + \lambda$  这样的  $w$  取 min，而不带负号 sqrt 并非下凸而是上凸。当然这里还是只需要考虑  $\sqrt{(j-i)(pre_j - pre_i)}$  是否满足四边形不等式。直观地理解的话，把  $pre_j$  替换成  $j$ ，于是  $w$  刚好到  $j-i$ ，一个凸性的临界点。然后  $pre$  可能会再稍微下凸一点，即考虑到  $\Delta pre$  单调不减，就让这个  $w$  变成下凸的了。

然后就是一个典型的区间分拆问题，这类问题听说只要把四边形不等式证出来了，则可推出 Monge 最短路关于经过的边数  $k$ （即拆分组数）有下凸性。这题貌似没有说清楚能不能让  $num_i = 0$ ，即某个区间为空，即只需要区间数  $\leq m$  而非恰好  $m$ 。但看起来通过一点调整可以说明组数越多越好。所以关于组数应该是一个下凸的减函数。

```
#include<bits/stdc++.h>
using namespace std;
const int N = 2e5 + 5;
int n, m;
int a[N];
long double pre[N];
int cnt[N];
long double dp[N];
long double w(int i, int j, long double lamb)
{
    return dp[i-1] + sqrtl( ( pre[j] - pre[i-1] )*(j-i+1) ) + lamb;
}
bool check(long double lamb)
{
    deque<pair<int, pair<int, int> > > q;
    dp[0] = 0;
    cnt[0] = 0;
    for(int i=1; i<=n; i++)
    {
        if(!q.empty() && q.front().second.second < i) q.pop_front();
        if(!q.empty() && q.front().second.first < i) q.front().second.first = i;
```

```

while(!q.empty() &&
    w(i, q.back().second.first, lamb)
    <= w(q.back().first, q.back().second.first, lamb) ) q.pop_back();
if(q.empty()) q.push_back({i, {i, n}});
else if( w(i, n, lamb) < w(q.back().first, n, lamb) )
{
    int l = q.back().second.first, r = n, ans = n;
    while(l<=r)
    {
        int mid = (l+r)/2;
        if( w(i, mid, lamb) < w(q.back().first, mid, lamb) )
        {
            ans = mid;
            r = mid-1;
        }
        else l = mid+1;
    }
    q.back().second.second = ans-1;
    q.push_back({i, {ans, n}});
}
dp[i] = w(q.front().first, i, lamb);
cnt[i] = cnt[q.front().first-1] +1;
}
return cnt[n] <= m;
}
void solve()
{
    cin>>n>>m;
    for(int i=1; i<=n; i++) cin>>a[i];
    sort(a+1, a+n+1);
    for(int i=1; i<=n; i++) pre[i] = pre[i-1] + a[i];

    long double l = 0, r = 1e14;
    for(int itr = 0; itr < 200; itr++)
    {
        long double mid = (l+r)/2;
        if( check(mid) )
        {
            r = mid;
        }
        else l = mid;
    }
    check(r);
    cout<<dp[n]-m*r<<"\n";
    return;
}
int main()
{

```

```

ios::sync_with_stdio(0);
cin.tie(0);
cout.tie(0);
cout<<fixed<<setprecision(15);
int T = 1;
// cin>>T;
while(T-->0)
{
    solve();
}
return 0;
}

```

## 2.2 维护凸包相关数据结构

### 2.2.1 李超线段树

值域  $10^6$ ，针对线段

要求在平面直角坐标系下维护两个操作：

- 在平面上加入一条线段。记第  $i$  条被插入的线段的标号为  $i$ 。
- 给定一个数  $k$ ，询问与直线  $x = k$  相交的线段中，交点纵坐标最大的线段的编号。

```

#include<bits/stdc++.h>
using namespace std;

class LiChaoTree
{
private:
    const double eps = 1e-9;
    int cmp(double x, double y)
    {
        if(x - y > eps) return 1;
        if(y - x > eps) return -1;
        return 0;
    }
    int RANGE;
    struct Line
    {
        double k, b;
    };
    vector<Line> line;
    vector<int> lazytag;
    double calc(int id, int x)
    {

```

```

        return line[id].b + line[id].k * x;
    }
void upd(int root, int cl, int cr, int u)
{
    int &v = lazytag[root];
    if(v == -1)
    {
        v = u;
        return;
    }
    int mid = (cl + cr) >> 1;
    int bmid = cmp(calc(u, mid), calc(v, mid));
    if( bmid == 1 || (!bmid && u < v) ) swap(u, v);
    int bl = cmp(calc(u, cl), calc(v, cl));
    if(bl == 1 || (bl==0 && u < v)) upd(root<<1, cl, mid, u);
    int br = cmp(calc(u, cr), calc(v, cr));
    if(br == 1 || (br==0 && u < v)) upd(root<<1|1, mid+1, cr, u);
    return;
}
void update(int root, int cl, int cr, int l, int r, int u)
{
    if(l <= cl && cr <= r)
    {
        upd(root, cl, cr, u);
        return;
    }
    int mid = (cl + cr)>>1;
    if( l <= mid) update(root<<1, cl, mid, l, r, u);
    if( mid+1 <= r) update(root<<1|1, mid+1, cr, l, r, u);
    return;
}
using pdi = pair<double, int>;
pdi pmax(pdi x, pdi y)
{
    if( cmp(x.first, y.first) == 1 ) return x;
    if( cmp(x.first, y.first) == -1 ) return y;
    return x.second < y.second ? x : y;
}
public:
LiChaoTree(int _RANGE)
{
    RANGE = _RANGE;
    lazytag = vector<int>(RANGE * 4 + 5, -1);
    return;
}
pdi query(int root, int l, int r, int x)
{
    if(r < x || x < l) return {0, 0}; // {-inf, id}

```

```

    int mid = (l+r)>>1;
    pdi res = {0, 0};
    if(lazytag[root] != -1) res = {calc(lazytag[root], x), lazytag[root]};
    if( l==r ) return res;
    return pmax(res, pmax(query(root<<1, l, mid, x), query(root<<1|1, mid+1, r, x)) );
}

void add(int x_0, int y_0, int x_1, int y_1)
{
    Line tmp;
    if(x_0 == x_1) tmp.k = 0, tmp.b = max(y_0, y_1); // only serve this problem
    else tmp.k = 1.0 * (y_1 - y_0) / (x_1 - x_0), tmp.b = y_0 - tmp.k * x_0;
    int id = line.size();
    line.push_back(tmp);
    update(1, 1, RANGE, x_0, x_1, id);
    return;
}

};
// only serve the problem https://www.luogu.com.cn/problem/P4097
const int MOD1 = 39989; // also serve as the range of x
const int MOD2 = 1000'000'000;
void solve()
{
    LiChaoTree lichao(MOD1);

    int q;
    cin>>q;
    lichao.add(1, 0, MOD1, 0); // only serve this problem
    for(int lastans = 0; q-->0; )
    {
        int op;
        cin>>op;
        if(op==0)
        {
            int x;
            cin>>x;
            x = (x + lastans - 1 + MOD1) % MOD1 + 1;
            cout<< (lastans = lichao.query(1, 1, MOD1, x).second) << "\n";
        }
        else if(op==1)
        {
            int x_0, y_0, x_1, y_1;
            cin>>x_0>>y_0>>x_1>>y_1;
            x_0 = (x_0 + lastans - 1 + MOD1) % MOD1 + 1;
            y_0 = (y_0 + lastans - 1 + MOD2) % MOD2 + 1;
            x_1 = (x_1 + lastans - 1 + MOD1) % MOD1 + 1;
            y_1 = (y_1 + lastans - 1 + MOD2) % MOD2 + 1;
            if(x_0 > x_1) swap(x_0, x_1), swap(y_0, y_1);
            lichao.add(x_0, y_0, x_1, y_1);
        }
    }
}

```



```

    }
}
return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0); cout.tie(0);
    int T = 1;
    // cin>>T;
    while(T--)
    {
        solve();
    }
    return 0;
}

```

这个模板其实可能不够好，也不常用。主要是它的时间复杂度会和定义域大小相关。有些问题要查  $1e9$  就寄了。

### 值域 $10^9$ ，针对直线

更常用的应该还是没有定义域限制，功能一样，然后拿来优化一些 dp。

下面这个模板展示了原来的 dp 和用李超线段树优化后的用法。

CF2107F2

```

#include<bits/stdc++.h>
using namespace std;

using ll = long long;
class LiChaoTree // get min on [0, 1e9]
{
private:
    static const int inf = 1e9, M = 2e5;
    class line
    {
    public:
        ll k, b;
        line(ll kk = 0, ll bb = 0) { k = kk, b = bb; }
        ll get(ll x) { return x * k + b; }
    };
    ll inter(line aa, line bb)
    {
        return (bb.b - aa.b) / (aa.k - bb.k);
    } // we only need the integer part, and (aa.k - bb.k) won't be 0 when we use it
    vector<pair<int, int>> ch;
    vector<line> seg;
    // only 1 new node for 1 add operation, so the memory is N instead of N log V
    int tot = 0; // avoid using .size() then .push_back(), because here we are using & for a vector object

```

```

// when .push_back(), the still using & for the object may be invalid
void _add(int &id, int L, int R, line now)
{
    if(!id) return id = ++tot, seg[id] = now, void();
    int mid = (L+R) >> 1;
    ll La = seg[id].get(L), Ra = seg[id].get(R), Lb = now.get(L), Rb = now.get(R);
    if(La <= Lb && Ra <= Rb) return;
    if(La >= Lb && Ra >= Rb) return seg[id] = now, void();
    ll it = inter(seg[id], now);
    if(La <= Lb)
    {
        if(it<=mid) _add(ch[id].first, L, mid, seg[id]), seg[id] = now;
        else _add(ch[id].second, mid + 1, R, now);
    }
    else
    {
        if(it <= mid) _add(ch[id].first, L, mid, now);
        else _add(ch[id].second, mid + 1, R, seg[id]), seg[id] = now;
    }
}

ll _qry(int id, int L, int R, int x)
{
    if(!id) return (ll)inf*inf;
    int mid = (L+R) >> 1;
    if(x <= mid) return min(seg[id].get(x), _qry(ch[id].first, L, mid, x));
    else return min(seg[id].get(x), _qry(ch[id].second, mid + 1, R, x));
}

public:
    int rt = 0;
    void add(line now) { return _add(rt, 0, inf, now), void(); }
    ll qry(int x) { return _qry(rt, 0, inf, x); }
    LiChaoTree(int _M = M) { seg = vector<line>(_M+5); ch = vector<pair<int, int>>(_M+5); }
};

const int N = 1e6 + 5;
int n;
long long a[N], dp[N], best_k_for_fixed_j[N];
void solve()
{
    cin>>n;
    for(int i=1; i<=n; i++) cin>>a[i];
    LiChaoTree lc_for_k(n), lc_for_dp(n);
    for(int i=1; i<=n; i++)
    {
        // for(int k=0; k<i; k++)
        // {
        // if(k==0) best_k_for_fixed_j[i] = - a[i] * k - k + dp[k];
        // else best_k_for_fixed_j[i] = min(best_k_for_fixed_j[i], - a[i] * k - k + dp[k]);
        // }
    }
}

```

```

        lc_for_k.add({-(i-1), -(i-1) + dp[i-1]});
        best_k_for_fixed_j[i] = lc_for_k.qry(a[i]);
        // dp[i] = dp[i-1] + a[i];
        // for(int j = 1; j<=i; j++) dp[i] = min(dp[i], (i-j) + a[j]*i + (i-1) + best_k_for_fixed_j[j] );
        lc_for_dp.add({ 1+a[i]+1, -i-1+best_k_for_fixed_j[i] });
        dp[i] = lc_for_dp.qry(i);
        cout<<dp[i]<<" ";
    }
    cout<<"\n";
    // cout<<dp[n]<<"\n";
    return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0); cout.tie(0);
    int T = 1;
    cin>>T;
    while(T--) solve();
    return 0;
}
/*
*/
*/

```

有的人会误认为内存要开成  $N \log V$ ，其中  $V$  为  $10^9$ ，而大多数情况下存一个直线是需要至少一个 `int` 一个 `long long` 的，如果  $N$  是  $10^6$  再乘个 30 就很容易炸。实际上这是一种误解。

仔细观察实现过程，这个线段树没有分割后分别递归，而是只会往一侧走。

每次存下直线后都会直接停止，从而可以知道每次 `add` 操作只会新建一个结点，所以内存开成  $N$  即可。

另外还有一个经典问题

Q: 这种问题可能是什么原因，孩子调半天没搞明白

- 想用 `vector` 动态开点，每次用 `.size()` 来得到当前的 `id`（本地测试能出正确答案，OJ 上 RE）
- 也是 `vector`，但是定义的时候预先定义大小，另外开一个 `int` 类型来记录当前 `id`（本地测出来和 1 一样，OJ 上 AC）

A: 递归中存在对 `vector` 的引用，结果中途给 `vector` 扩容了，导致引用失效。

### 若干个树共享内存池

有些题目得不同组分别维护李超树，也不太能事先知道每棵加进的直线数，只保证加进的直线数总和，前面又说了目前这个代码不适合做 `vector` 扩容，那么目前只能共享一下内存池。由于我第一次知道 Cpp 的静态变量和 Java 不一样，类内声明完后还得在外面声明一次，所以目前的模板还有点臃肿。

需要用到这个写法的题目其它部分太长，只贴一个不需要这个写法的简单题的应用

```

#include<bits/stdc++.h>
using namespace std;

```

```

using ll = long long;
class LiChaoTree // get min on [0, 1e9]
{
private:
    class line
    {
    public:
        ll k, b;
        line(ll kk = 0, ll bb = 0) { k = kk, b = bb; }
        ll get(ll x) { return x * k + b; }
    };
    static const int inf = 1e9;
    static vector<pair<int, int> > ch;
    static vector<line> seg;
    static int tot;

    ll inter(line aa, line bb)
    {
        return (bb.b - aa.b) / (aa.k - bb.k);
    } // we only need the integer part, and (aa.k - bb.k) won't be 0 when we use it
    void _add(int &id, int L, int R, line now)
    {
        if(!id) return id = ++tot, seg[id] = now, void();
        int mid = (L+R) >> 1;
        ll La = seg[id].get(L), Ra = seg[id].get(R), Lb = now.get(L), Rb = now.get(R);
        if(La <= Lb && Ra <= Rb) return;
        if(La >= Lb && Ra >= Rb) return seg[id] = now, void();
        ll it = inter(seg[id], now);
        if(La <= Lb)
        {
            if(it <= mid) _add(ch[id].first, L, mid, seg[id]), seg[id] = now;
            else _add(ch[id].second, mid + 1, R, now);
        }
        else
        {
            if(it <= mid) _add(ch[id].first, L, mid, now);
            else _add(ch[id].second, mid + 1, R, seg[id]), seg[id] = now;
        }
    }
    ll _qry(int id, int L, int R, int x)
    {
        if(!id) return (ll)inf*inf;
        int mid = (L+R) >> 1;
        if(x <= mid) return min(seg[id].get(x), _qry(ch[id].first, L, mid, x));
        else return min(seg[id].get(x), _qry(ch[id].second, mid + 1, R, x));
    }
public:
    int rt = 0;

```

```
void add(line now) { return _add(rt, 0, inf, now), void(); }
ll qry(int x) { return _qry(rt, 0, inf, x); }
};
const int N = 2e5 + 5;
vector<pair<int, int>> LiChaoTree::ch = vector<pair<int, int>>(N);
vector<LiChaoTree::line> LiChaoTree::seg = vector<line>(N);
int LiChaoTree::tot = 0;

ll B[N], C[N];
void solve()
{
    int n, m;
    cin >> n >> m;
    for(int i=1; i<=n; i++) cin >> B[i];
    sort(B+1, B+n+1);
    LiChaoTree lc;
    for(int i=1; i<=n; i++) lc.add({ -(n+1-i), -B[i] * (n+1-i) }); // use - to solve max
    for(int i=1; i<=m; i++)
    {
        cin >> C[i];
        cout << -lc.qry(C[i]) << " ";
    }
    return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0); cout.tie(0);
    int T = 1;
    // cin >> T;
    while(T--) solve();
    return 0;
}
```

## 第 3 章 计算几何

### 3.1 随机化方法

#### 3.1.1 最小公共圆

一个期望复杂度正确的随机化算法

```
#include<bits/stdc++.h>
using namespace std;
const double eps = 1e-9;
pair<double, double> getCircle(pair<double, double> p1, pair<double, double> p2, pair<double, double> p3) {
    double A = p2.first - p1.first;
    double B = p2.second - p1.second;
    double C = p3.first - p1.first;
    double D = p3.second - p1.second;
    double E = A * (p1.first + p2.first) + B * (p1.second + p2.second);
    double F = C * (p1.first + p3.first) + D * (p1.second + p3.second);
    double det = 2 * (A * D - B * C);
    if (abs(det) < eps)
    {
        return {0, 0};
    }
    double centerX = (D * E - B * F) / det;
    double centerY = (A * F - C * E) / det;
    return {centerX, centerY};
}
void solve()
{
    int n;
    cin>>n;
    vector<pair<double, double> > p(n);
    for(int i=0; i<n; i++) cin>>p[i].first>>p[i].second;
    random_shuffle(p.begin(), p.end());
    auto dis = [&](pair<double, double> x, pair<double, double> y)
    {
        return sqrt( pow(x.first - y.first, 2) + pow(x.second - y.second, 2) );
    };
    pair<double, double> O = p[0];
    double R = 0;
```

```
for(int i=1; i<n; i++)
{
    if(dis(p[i], 0) > R + eps)
    {
        O = p[i], R = 0;
        for(int j=0; j<i; j++)
        {
            if(dis(p[j], 0) > R + eps)
            {
                O.first = (p[i].first + p[j].first)/2;
                O.second = (p[i].second + p[j].second)/2;
                R = dis(p[j], 0);
                for(int k=0; k<j; k++)
                {
                    if(dis(p[k], 0) > R + eps)
                    {
                        O = getCircle(p[i], p[j], p[k]);
                        R = dis(p[k], 0);
                    }
                }
            }
        }
    }
}
cout<<R<<"\n"<<O.first<<" "<<O.second<<"\n";
return;
}
int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0);
    cout.tie(0);
    cout<<fixed<<setprecision(10);
    int T = 1;
    // cin>>T;
    while(T--)
    {
        solve();
    }
    return 0;
}
```

## 3.2 数值算法

### 3.2.1 自适应 Simpson 积分

```
#include<bits/stdc++.h>
using namespace std;
using ll = long long;
const int N = 5e5 + 5;
const double EPS = 1e-7;

double a, b, c, d;
double f(double x)
{
    return (c*x+d) / (a*x+b);
}
double simpson(double l, double r)
{
    double mid = (l+r)/2;
    return (r-l) * (f(l) + 4 * f(mid) + f(r)) /6;
}
double asr(double l, double r, double ans, double eps, int step)
{
    {
        double mid = (l+r)/2;
        double fl = simpson(l, mid), fr = simpson(mid, r);
        if( abs(fl + fr - ans) <= 15 * eps && step < 0 )
        {
            return fl + fr + (fl + fr - ans) / 15;
        }
        return asr(l, mid, fl, eps/2, step-1) + asr(mid, r, fr, eps/2, step-1) ;
    }
}
double calc(double l, double r)
{
    return asr(l, r, simpson(l, r), EPS, 12);
}

void solve()
{
    int L, R;
    cin>>a>>b>>c>>d>>L>>R;
    cout<<fixed<<setprecision(6)<<calc(L, R)<<"\n";
    return;
}

int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0), cout.tie(0);

    int T=1;
    // cin>>T;
    while(T--) solve();
}
```



```
    return 0;  
}
```

## 第 4 章 数学

### 4.1 初等数论相关

#### 4.1.1 找模素数 $P$ 的原根

首先必然素数  $P$  必然有原根。据说可以证明最小的原根的数量级大概是  $O(P^{1/4})$ ，然后考虑到 Fermat 小定理给出的  $p-1$  必然是阶的倍数，所以可以暴力枚举其因数判断阶是否小于  $p-1$ 。

```
#include<bits/stdc++.h>
using namespace std;
class FindPrimitiveRoot
{
private:
    int P;
    long long power_mod_P(long long x, int idx)
    {
        long long res = 1;
        for(long long now = x; idx>0; now = now*now%P, idx>>=1)
        {
            if(idx & 1) res = res*now%P;
        }
        return res;
    }
public:
    FindPrimitiveRoot(int _P)
    {
        P = _P;
        return;
    }
    int work()
    {
        if(P==2) return 1;
        vector<int> d;
        for(int i=1; i*i<=(P-1); i++)
        {
            if( (P-1)%i !=0 ) continue;
            d.push_back(i);
            if(i!=1 && i*i!=(P-1) ) d.push_back( (P-1)/i );
        }
        // the possible choices to be period must be the divisor of (P-1)
    }
};
```

```

    for(int i=2; i<P; i++)
    {
        bool flag = 1;
        for(auto j : d)
        {
            if(power_mod_P(i, j) == 1)
            {
                flag = 0;
                break;
            } // P-1 isn't the least period
        }
        if(flag) return i;
    } // for prime, the least root is small, so just check from 2
    return -1; // for legal case, it's not gonna happen
}
};
int main()
{
    return 0;
}

```

### 4.1.2 杜教筛求 premu

```

#include<bits/stdc++.h>
using namespace std;
const int M = 2e7+5;
bool vis[M];
int mu[M], premu[M];
vector<int> prime;
map<long long, long long> S0;
long long sum(long long n)
{
    if(n<M) return premu[n];
    if(S0.count(n)) return S0[n];
    long long res = 1;
    for(long long t=2, r; t<=n; t = r+1)
    {
        r = min(n/(n/t), n);
        res -= sum(n/t)*(r-t+1);
    }
    S0[n] = res;
    return res;
}
void sieve()
{
    premu[1] = mu[1] = 1;
    for(int j=2; j<M; j++)

```

```

{
    if(!vis[j])
    {
        prime.push_back(j);
        mu[j] = -1;
    }
    premu[j] = premu[j-1] + mu[j];
    for(auto p : prime)
    {
        if(1ll*p*j<M)
        {
            vis[p*j] = 1;
            if(j%p == 0)
            {
                mu[p*j] = 0;
                break;
            }
            else mu[p*j] = -mu[j];
        }
        else break;
    }
}
return;
}
int main()
{
    return 0;
}

```

## 4.2 多项式相关

### 4.2.1 NTT 以及半在线卷积

主要的部分就是算一个特殊的矩阵乘向量得到的向量 void ntt(vector a, bool invert 因为逆矩阵几乎一样, 方法通用, 就写进一个函数里) 记录一下 a 的大小 n。实际上这里的 n 会在后面调用的时候提前变成 2 的幂次。

先作一个预处理, 所谓蝴蝶操作。就是后面我们会把系数根据奇偶位分成两份多项式, 直到拆到底层成为一个直接可以引用的下标。我们把这个下标提前放好, 后面就可以迭代地自下往上了。这里先直观给出小例子

```

0 → 000 → 000 → 0
1 → 001 → 100 → 4
2 → 010 → 010 → 2
3 → 011 → 110 → 6
4 → 100 → 001 → 1
5 → 101 → 101 → 5

```

$6 \rightarrow 110 \rightarrow 011 \rightarrow 3$

$7 \rightarrow 111 \rightarrow 111 \rightarrow 7$

大致上是一个两两配对并 swap 的过程。我们可以只枚举小的然后去找对应大的，防止 swap 两次结果还原了。只要能做到  $O(n \log n)$  就行了，这里的实现方法可以是：考虑我给  $i$  从 0 开始加实际上在干什么：从右往左找到第一个 0，变成 1，然后抹掉这个位置以右的位置。镜面地对另一个  $j$  从 0 开始操作，就能得到镜面的结果了。

从最短  $\text{len}=2$  开始往  $n$  合并，每次  $\text{len}$  倍长。假设已经计算出了  $\text{len}/2$  的奇数项多项式和偶数项多项式。现在来计算  $\text{len}$  的多项式。记当前单位根是  $w_{\text{len}}$ ，实质上对于  $f(w_{\text{len}}^k)$ ，我们需要的是  $f(w_{\text{len}}^k) = f_1(w_{\text{len}/2}^k) + w_{\text{len}}^k f_2(w_{\text{len}/2}^k)$ 。对于  $k < \text{len}/2$  我们就直接引用对应位置的点值来计算；对于  $k \geq \text{len}/2$  的，考虑到  $w_{\text{len}/2}^k = w_{\text{len}/2}^{k-\text{len}/2}$ ， $w_{\text{len}}^k = -w_{\text{len}}^{k-\text{len}/2}$ ，可以引用对应点值，也可以在处理  $k - \text{len}/2$  时顺便处理了，并且这样的话就可以直接覆盖掉对应的位置而不怕后面还会被引用了。

模意义下的单位根  $w_{\text{len}}$  是什么意思呢？NTT 模数  $p$  减一，即 998244353-1 后是  $2^{23} \approx 8e6$  的倍数，大于算法竞赛环境下  $O(n \log n)$  复杂度允许的  $n$ ；3 又是这个模数的原根，即第一个循环节刚好是费马小定理指出的一个循环节  $p-1$ ，所以我们可以指定  $w_{\text{len}} = 3^{(p-1)/\text{len}}$ ，并验证它符合我们需要的各种单位根性质。稍微强调一下：最主要的动机在于

$$\sum_{k=0}^{n-1} (w_n^{i-j})^k$$

应该需要在  $i = j$  时为一个统一的常数，而在  $i \neq j$  时为 0，即这时

$$w_n^{i-j} \neq 1 \wedge w_n^{(i-j)n} = 1$$

加入让  $w_{\text{len}} = g^{(p-1)/\text{len}}$ ，其中  $g$  不是原根的话， $g^{(i-j)(p-1)/n}$ ，这里  $(i-j)$  的取值范围实际上是  $-(n-1)$  到  $(n-1)$ ，就有可能撞上一个比  $p-1$  小的循环节，让这个结果变成 1 了。这个动机实质上就是在验证这个形式的矩阵是否真的是逆矩阵。对于原根打底的，那确实是。反之，要么我们通过上面的解剖来验证伪，要么就直接考虑一下矩阵中存在完全相同的某两行，那么行列式为 0，必然不存在逆矩阵。

如果是 invert 的话实质上求逆时还会多一项公因数  $1/n$ （其实就是我们先前提到的统一常数），所以每项结果再乘一下就好了。

下面顺便把半在线卷积（一个类似 CDQ 分治的东西），以及可以利用半在线卷积实现的功能写上去了。实际上现场推也并不困难。

```
#include<bits/stdc++.h>
using namespace std;

using ll = long long;

const int MOD = 998244353, G = 3;

ll power(ll x, ll y)
{
    ll res = 1;
    for( ; y>0; y>>=1, x=x*x%MOD) if(y&1) res = res*x%MOD;
    return res;
}

ll inv(ll x)
{

```

```

        return power(x, MOD-2);
    }

class NTT
{
private:
    static void ntt(vector<int> &a, bool invert)
    {
        int len = a.size();
        for(int now = 0, i = 1; i<len-1; i++)
        {
            int hbit = len/2;
            while( (now & hbit) )
            {
                now ^= hbit;
                hbit>>=1;
            }
            now^=hbit;
            if(now > i) swap(a[now], a[i]);
        }
        for(int l = 2; l<=len; l<<=1)
        {
            int w1 = power(G, (MOD-1)/l);
            if(invert) w1 = power(w1, MOD-2);
            for(int i=0; i<len; i+=l)
            {
                for(int k=0, now = 1; k<l/2; k++, now = 1ll*now*w1%MOD)
                {
                    int u = a[i+k], v = 1ll*a[i+k+l/2]*now%MOD;
                    a[i+k] = (u+v)%MOD;
                    a[i+k+l/2] = (u-v+MOD)%MOD;
                }
            }
        }

        if(invert)
        {
            int invn = power(len, MOD-2);
            for(int i=0; i<len; i++) a[i] = 1ll*invn*a[i]%MOD;
        }

        return;
    }

public:
    static vector<int> multiply(vector<int> vec1, vector<int> vec2)
    {
        vector<int> res;
    }

```

```

        int len = 1;
        while(len < (int)(vec1.size() + vec2.size()-1) ) len<=1;
        vec1.resize(len, 0);
        vec2.resize(len, 0);
        ntt(vec1, 0);
        ntt(vec2, 0);
        for(int i=0; i<len; i++)
        {
            res.push_back(1ll*vec1[i]*vec2[i]%MOD);
        }
        ntt(res, 1);
        return res;
    }
};

class Conv
{
private:
    void divide_solve(int l, int r)
    {
        if(l==r)
        {
            func(f[l], l);
            return;
        }
        int mid = (l+r)/2;
        divide_solve(l, mid);
        vector<int> tf(mid-l+1), tg(r-l+1);
        for(int i=l; i<=mid; i++) tf[i-l] = f[i];
        for(int i=1; i<=r-l; i++) tg[i] = g[i];
        auto res = NTT::multiply(tf, tg);
        for(int i=mid+1; i<=r; i++)
        {
            f[i] = ( f[i] + res[i-l] ) %MOD;
        }
        divide_solve(mid+1, r);
        return;
    }
public:
    vector<int> f, g;
    function<void(int &, int)> func;
    void self_convolution()
    {
        int len = 1;
        while(len < (int)g.size()) len<=1;
        g.resize(len);
        f = vector<int>(len, 0);
        f[0] = 1;
    }
};

```

```

        divide_solve(0, len-1);
        return;
    }
};

class Poly
{
public:
    static vector<int> exp(vector<int> G)
    { // guarantee G[0] = 0;
        int len = G.size();
        for(int i=1; i<len; i++) G[i] = 1ll * G[i] * i %MOD;
        Conv T;
        T.func = [&](int &x, int y)
        {
            if(y!=0) x = 1ll * x * inv(y) %MOD;
            return;
        };
        T.g = G;
        T.self_convolution();
        return T.f;
    }
    static vector<int> poly_inv(vector<int> G)
    { // guarantee G[0] = 1
        Conv T;
        T.func = [&](int &x, int y)
        {
            if(y!=0) x = ( MOD - x ) %MOD;
            return;
        };
        T.g = G;
        T.self_convolution();
        return T.f;
    }
    static vector<int> ln(vector<int> G)
    { // guarantee G[0] = 1
        vector<int> G_inv = poly_inv(G);
        int len = G.size();
        vector<int> G_p(len);
        for(int i=1; i<len; i++) G_p[i-1] = 1ll * G[i] * i %MOD;
        auto res = NTT::multiply(G_p, G_inv);
        vector<int> F(len);
        F[0] = 0;
        for(int i=1; i<len; i++) F[i] = 1ll * res[i-1] * inv(i) %MOD;
        return F;
    }
};

```



```
void init()
{
    return;
}

void solve()
{
    int n;
    cin>>n;
    vector<int> f(n);
    for(int i=0; i<n; i++) cin>>f[i];
    auto res = Poly::ln(f);
    for(int i=0; i<n; i++) cout<<res[i]<<"_";
    cout<<"\n";

    return;
}
/*

*/

int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0); cout.tie(0);

    // freopen("poly.in", "r", stdin);

    int T = 1;
    // cin>>T;
    while(T--) solve();
    return 0;
}
/*

*/
```

## 第 5 章 字符串

### 5.1 自动机

#### 5.1.1 AC 自动机

```
#include<bits/stdc++.h>
using namespace std;

using ll = long long;
const int N = 2e5 + 5, SIZE = 2e5 + 5;

class AC
{
public:
    struct Node
    {
        int son[26];
        int ans, fail, dg, idx;
        Node()
        {
            memset(son, 0, sizeof(son));
            ans = fail = idx = 0;
        }
    } tr[SIZE];

    int tot, ans[N], pid;
    AC()
    {
        tot = 0; pid = 0;
    }
    void insert(string s, int &idx)
    {
        int u = 0;
        int len = s.size();
        for(int i=0; i<len; i++)
        {
            int &son = tr[u].son[s[i] - 'a'];
            if(!son) son = ++tot;
```

```
        u = son;
    }
    if(!tr[u].idx) tr[u].idx = ++pidx;
    idx = tr[u].idx;
}
void build()
{
    queue<int> q;
    for(int i=0; i<26; i++)
    {
        if(tr[0].son[i]) q.push(tr[0].son[i]);
    }
    while(!q.empty())
    {
        int u = q.front();
        q.pop();
        for(int i=0; i<26; i++)
        {
            if(tr[u].son[i])
            {
                tr[tr[u].son[i]].fail = tr[tr[u].fail].son[i];
                tr[tr[tr[u].fail].son[i]].dg++;
                q.push(tr[u].son[i]);
            }
            else
            {
                tr[u].son[i] = tr[tr[u].fail].son[i];
            }
        }
    }
}
void query(string t)
{
    int u = 0;
    int len = t.size();
    for(int i=0; i<len; i++)
    {
        u = tr[u].son[t[i] - 'a'];
        tr[u].ans++;
    }
}
void topu()
{
    queue<int> q;
    for(int i=0; i<=tot; i++)
    {
        if(tr[i].dg == 0) q.push(i);
    }
}
```

```
        while(!q.empty())
        {
            int u = q.front();
            q.pop();
            ans[tr[u].idx] = tr[u].ans;
            int v = tr[u].fail;
            tr[v].ans += tr[u].ans;
            if(!--tr[v].dg) q.push(v);
        }
    }
};

int n;
int idx[N];

void solve()
{
    AC acm;
    cin>>n;
    string s;
    for(int i=1; i<=n; i++)
    {
        cin>>s;
        acm.insert(s, idx[i]);
        acm.ans[i] = 0;
    }
    acm.build();
    cin>>s;
    acm.query(s);
    acm.topu();
    for(int i=1; i<=n; i++) cout<<acm.ans[idx[i]]<<"\n";

    return;
}

int main()
{
    ios::sync_with_stdio(0);
    cin.tie(0), cout.tie(0);

    int T=1;
    // cin>>T;
    while(T--) solve();

    return 0;
}
```