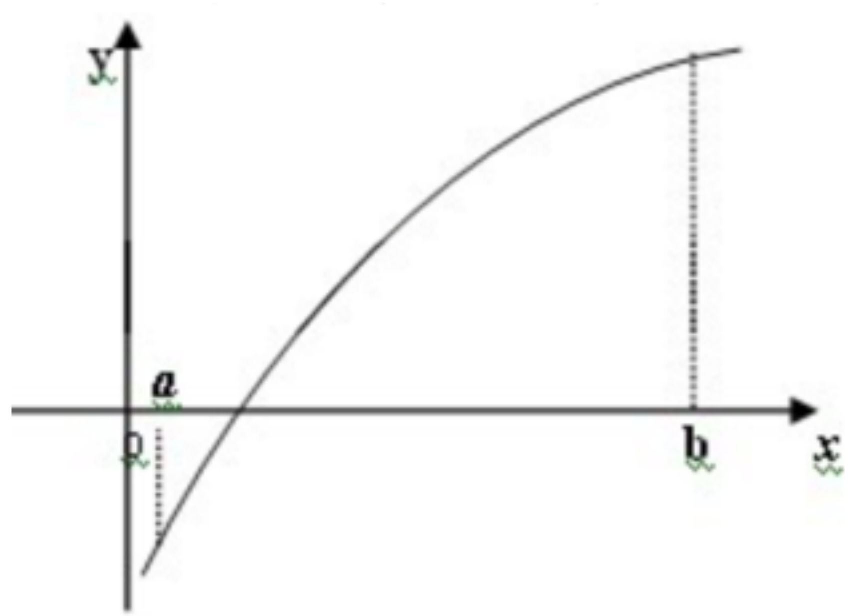
问求算法设计与实践01

二分,三分与倍增

二分

- 在一个**有序**序列中，查找一个元素
- 二分法求函数零点/方程的根



- hint:优选法

二分code

```
int t=-1,l=start,r=end,mid;
while (l<=r)
{
    mid=(l+r)>>1;
    if (Check(mid)) l=mid+1,t=mid;
    else r=mid-1;
}
```

二分code

```
int t=-1,l=start,r=end,mid;
while (l<=r)
{
    mid=(l+r)>>1;
    if (Check(mid)) l=mid+1,t=mid;
    else r=mid-1;
}
```

t=-1??? 连续???

Problem 3

- 给定一棵排序二叉树的先序遍历,求该树的后序遍历

problem 2

- 将 N 个人分成两组,其中 M 对人之间不和谐,即其中如果第 i 对的两个人在同一组, 那么就会有 Ci的不和谐值。
- 要求找出一个分组方案,使得最大不和谐值最小。
- $N, M \leq 10^6$

problem 2 cont’d

- 本题是一个要求 **最大值最小** 的优化问题
- 令F(x)为最大不和谐值不超过x时是否存在方案

problem 2 cont’d

- 本题是一个要求 **最大值最小** 的优化问题
- 令F(x)为最大不和谐值不超过x时是否存在方案
- F(x)=-1表示不存在
- F(x)=1表示存在
- F形如-1,-1,-1,-1,-1,1,1,1,1,1,1,1

problem 2　cont’d

- 二分x后，将图中边权小于x的边删去

problem 2 cont’d

- 二分x后，将图中边权小于x的边删去
- 问题转化为:
  - 对于一个给定的图，是否能将其分成两个部分，使得所有的边都是连接两个部分.
- 二分图判定

二分图判定

- 将第一个点放到左边，并加入队列
- 对队列中的当前点，遍历其所有边
- 每条边的另一端点有三种情况
  - 1，尚未入队：加入队列，并放到另一边
  - 2，已入队，不在同一边
  - 3，已入队，在同一边：不能形成二分图
- hint:一个图是否可二分和其是否有奇环等价，并查集实现?

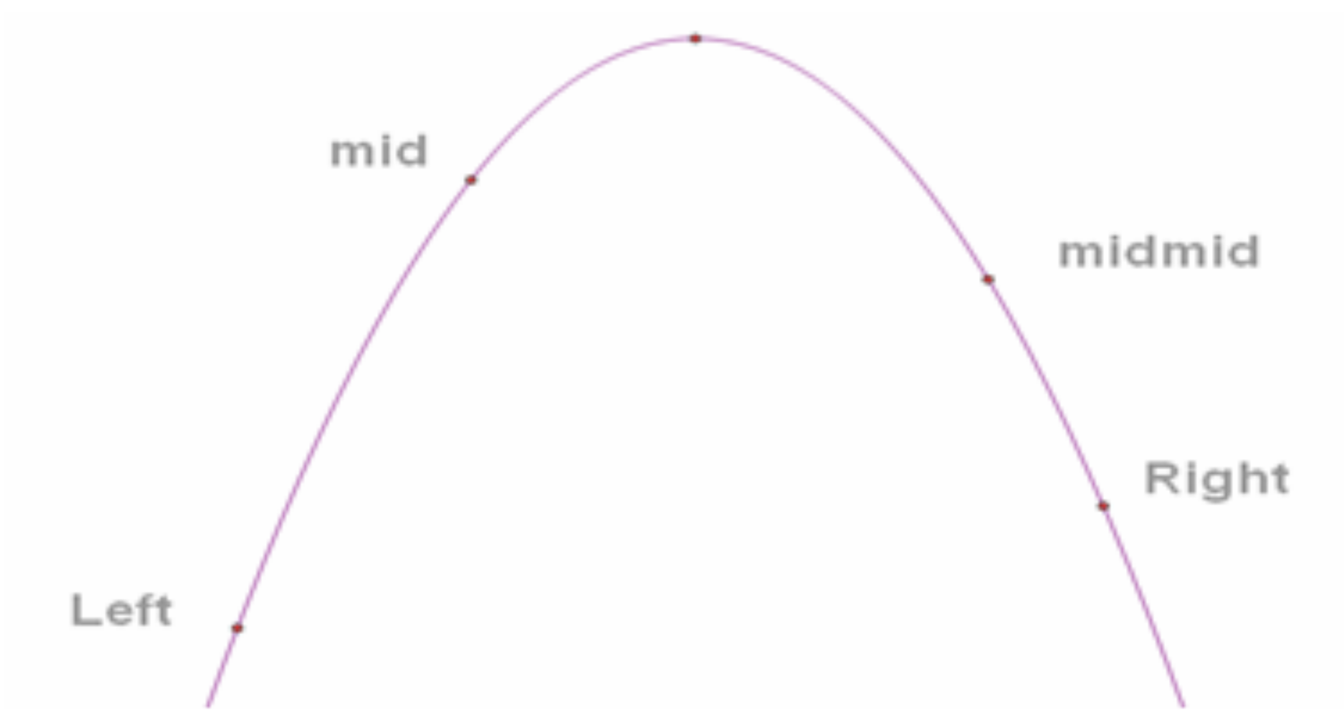
problem 2　cont'd

- 二分图判定的时间 O(N+M)
- 二分x的时间 O(log max(Ci))
- 二分是一种常用的将优化问题转化为判定问题的方法

当函数不再过零点

- -1,-1,-1,-1,-1,1,1,1,1的函数太“单调”，有时函数性质并没有这么好
- 如果是一个单峰函数呢?
- ex: f(x)=ax^2+bx+c

三分法



problem 6

- 已知两条线段 AB 和 CD. 一人在 AB 上的速度为 p,在 CD 上速度为 q, 否则速度为 r.
- 求从A到D的最小时间

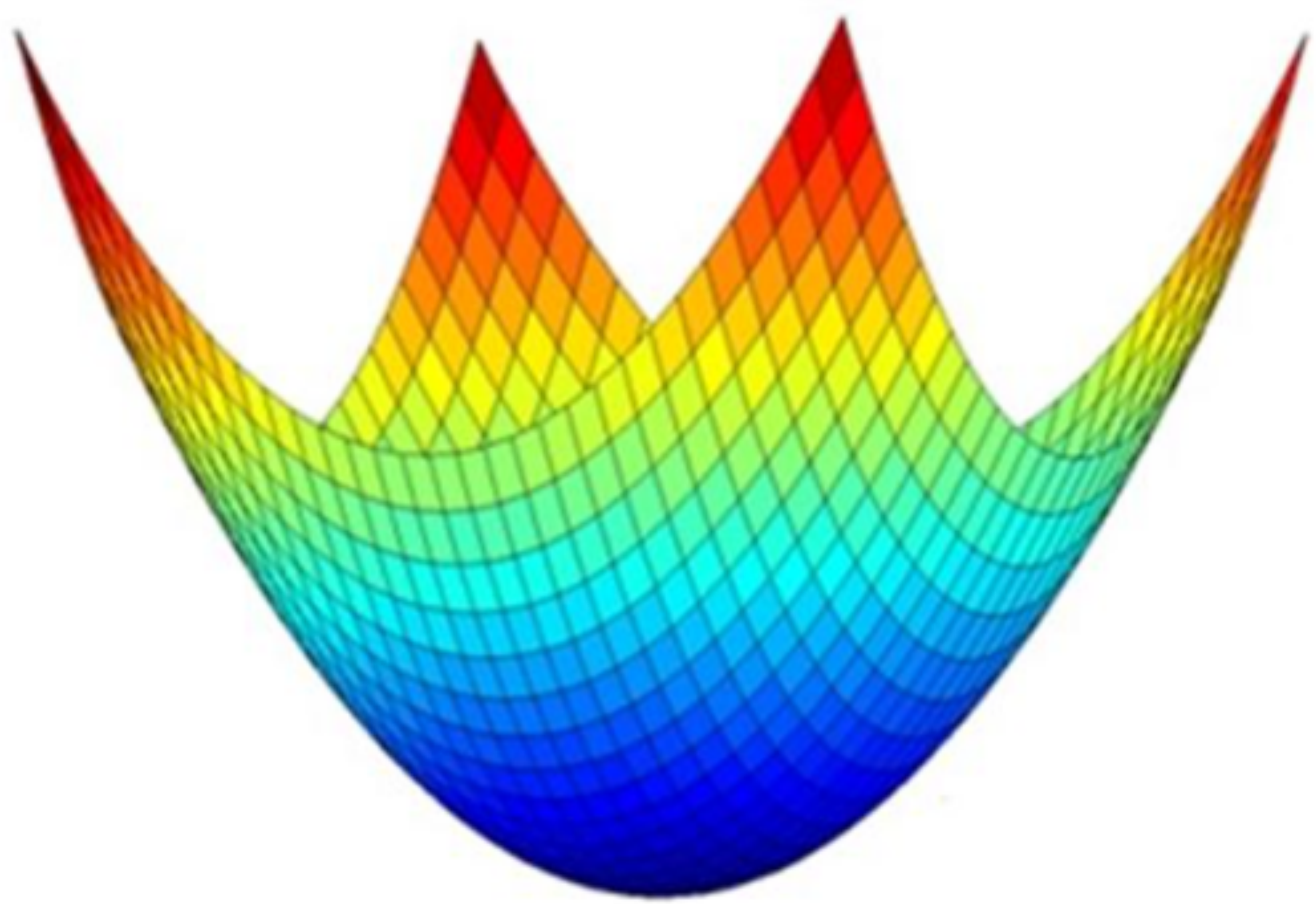
problem 6 cont’d

- 最优路径一定是在AB上走一段(可为0)，然后到CD上某点，再到D
- f(x), g(x)
- f(x),g(x)都是单峰函数

problem 6　cont’d

- 首先对f(x)三分
- 得到一个当前的midl或midr后，g(x)又是一个单峰函数
- 继续对g(x)三分
- 最后得到最优的f(x)

当函数不是一维的时候



爬山法

- 爬山法可对一个黑箱函数求出局部极值.

- 当函数并非单一极值时?

- hint:模拟退火，演化计算，量子退火
- https://en.wikipedia.org/wiki/Evolutionary_algorithm

problem 7

- 给定一个包含 N 个点的凸多边形,找出一个点使得其到 N 个顶点的距离之和最小.

problem 7   cont’d

- 可用爬山法，嵌套三分等方法解
- 如何确定唯一极值? 非凸多边形可以吗?
- 并非黑盒函数，是否有数学解析解?

- hint:凸函数,费马点

Optimization

- 现在人们已经能解决凸优化问题
  - 线性规划，二次规划，半正定规划…

- hint:《Convex Optimazation》, dual problem
- gradient descent, Conjugate gradient descent
- interior-point method

倍增算法

- 如何设计最少的面值拼出连续最大的面值?
- 答案自然是1,2,4,8..即2^k
- 倍增算法就基于这个性质，我们通过解决所有2^k的问题来解决整个问题

problem 1

Problem 1. 给定一个 N 个数的序列 S_i，要求回答 M 个询问. 每个询问以下两种之一

- 询问 $\sum_{i=a}^{b} S_i$
- 询问 $\max(S_a, S_{a+1}, ..., S_b)$

$N, M \leq 10^5$

problem 1　cont’d

- 询问sum
  - 由于没有修改，我们可用预处理出**前缀和**
- P[i]=S[1]+S[2]+…+S[i]

problem 1　cont’d

- 询问sum
  - 由于没有修改，我们可用预处理出**前缀和**
- P[i]=S[1]+S[2]+…+S[i]
- sum(S[a],S[a]+1,…S[b])=P[b]-P[a-1]

RMQ

- RMQ并不满足sum问题的性质
- 区间加法，**区间减法**
- 众数，中位数…

- hint:https://en.wikipedia.org/wiki/Range_minimum_query

倍增算法(Sparse Table)解RMQ

- 我们用F[i][k]表示从i开始，连续2^k个数的最小值
- 那么有F[i][0]=S[i]

倍增算法(Sparse Table)解RMQ

- 我们用F[i][k]表示从i开始，连续2^k个数的最小值
- 那么有F[i][0]=S[i]
- 2^k个数的最小值是 前2^(k-1)个数的最小值 或者后2^(k-1)个数的最小值。
- F[i][k]=min{F[i][k-1],F[i+2^(k-1)][k-1]}

倍增算法(Sparse Table) 解RMQ　cont’d

- 至于回答，类似拼钱, 二进制分解
- 如回答2~6的最小值，5=(101)2

倍增算法(Sparse Table)解RMQ　cont’d

- 至于回答，类似拼钱, 二进制分解
- 如回答2~6的最小值，5=(101)2
- 我们先找到2^2=4，用F[2][2]更新答案

倍增算法(Sparse Table)解RMQ　cont’d

- 至于回答，类似拼钱, 二进制分解
- 如回答2~6的最小值，5=(101)2
- 我们先找到2^2=4，用F[2][2]更新答案
- 那么现在变成6~6，我们找到2^0=1
- 用F[6][0]更新答案

倍增算法(Sparse Table)解RMQ　cont'd

- 至于回答，类似拼钱, 二进制分解
- 如回答2~6的最小值，5=(101)2
- 我们先找到2^2=4，用F[2][2]更新答案
- 那么现在变成6~6，我们找到2^0=1
- 用F[6][0]更新答案
- 预处理F需要O(NlogN)，每个回答需要O(logN)

problem 5

Problem 5. 求出斐波那契数列第 N 项对 10007 取模的结果.

$N <= 10^9$

problem 5   cont’d

- 通项公式，数学性质?

- 先来看下快速幂

快速幂

- 如何快速计算 6^66666666
- 假设乘法和加法运算只需常数时间
- 或者取模

快速幂　cont’d

- k^N
- a naive sol, O(N)

快速幂　cont’d

- k^N
- a naive sol, O(N)
- 类似刚才倍增的拼钱，将N二进制拆分
- $N=(110101100)_2$
- 求出k^1, k^2, k^4, k^8,…, k^[logN]

快速幂 cont’d

- k^N
- a naive sol, O(N)
- 类似刚才倍增的拼钱，将N二进制拆分
- N=$(110101100)_2$
- 求出k^1, k^2, k^4, k^8,…, k^[logN]
- 再有选择的取出若干项相乘即可
- 只需O(logN)次乘法

矩阵快速幂

- 比如求A^N, A是一个k*k的矩阵
- 我们至少可以在O(k^3 logN)的时间内求出

problem 5 cont'd

- 将递推关系改写成矩阵乘法形式

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a+b \end{pmatrix}$$

- 由矩阵乘法结合律，我们先计算第一个矩阵的N次方，再与第二个矩阵相乘即可
- 复杂度O(8*logN)
- hint:线性常系数齐次递推，生成函数
- http://www.matrix67.com/blog/archives/276

problem 5 cont'd

- $f(n)=f(n-1)+f(n-2)+n$

problem 8

Problem 8 (*). 在数轴上有 N 类点，每一类用一个三元组 (S, E, D) 来表示，意思是这些点分布在

$$S, S+D, S+2D, ..., S+kD \qquad (S+kD \leq E)$$

上. 已知最多只有一个坐标上有奇数个点，要求找出它或指出不存在。

$N <= 10^5$, $S, E, D <= 10^9$

problem 8　cont’d

- 我们用0表示偶数个点，1表示奇数个点
- 那么数轴上的点分布如下
- 00000000000000000001 0000..
- 因为数轴太大，点太多
- 我们无法快速判断1的位置

problem 8 cont’d

- 000000000000000000010000..
- 考虑前缀和
- 000000000000000000011111..
- 我们可以用二分法找到0/1分界点，即唯一的1的位置
- 每次二分时，扫描每一类点，统计点的前缀和
- 复杂度O(N*logS)，S为数轴长，即值域

problem 4

Problem 4. 如果在一个长度为 $2K$ 的序列中，前 K 个数的和小于 S，后 K 个数的和也小于 S，我们就称这个序列是有趣的。

现给定一个 N 个元素的序列，求以序列中每个元素为开端的连续最长有趣序列的长度.（即输出 N 个长度）

$N \leq 10^5$

problem 4   cont'd

- 枚举i，二分最远j使得i~j与j+1~j+j-i+1为有趣的
- i, … , j , j+1,…, j+j-i+1
- 时间复杂度O(NlogN)

problem 4　cont’d

- 枚举i，二分最远j使得i~j与j+1~j+j-i+1为有趣的
- i, … , j , j+1,…, j+j-i+1
- 时间复杂度O(NlogN)
- 看似没问题的二分算法其实是错误的
- 如S=100,N个数为
- 1 1 98 98 1 1
- 当i=1，二分j=2时不合法，而其实j=3时合法

problem 4   cont’d

- 考虑枚举起点再二分是不对的
- 但是如果枚举中间点再二分是没有错的
- 所以我们枚举中间点，再二分最远扩展距离
- 这样我们得到若干区间
- 再处理一下就能得到答案

More..

- LCA与RMQ
- 分治法解平面最近点对问题
- O(k^2logN)的快速幂
- …