算法题准备材料

Tyrael.wang

# JAVA常用API

整数

public static int bitCount(int i)

https://blog.csdn.net/zhouzipeng000/article/details/56676885

1. 栈Stack类继承Vector类。主要方法：pushpoppeekempty（Java编程思想：Java本身的栈设计欠佳。暴露了很多Vector的方法。）

2. 队列Java 有Queue接口。LinkedList实现了Queue接口。offer/add插入队尾remove、pollelement、peek3. 字符串相关StringBuffer 线程安全StringBuilder 非线程安全字符串逆转stringBuilder.reverse4. 集合Arrays.binSearch的返回值

**private static int** binarySearch0(**int**[] a, **int** fromIndex, **int** toIndex,  
 **int** key) {  
 **int** low = fromIndex;  
 **int** high = toIndex - 1;  
  
 **while** (low <= high) {  
 **int** mid = (low + high) >>> 1;  
 **int** midVal = a[mid];  
  
 **if** (midVal < key)  
 low = mid + 1;  
 **else if** (midVal > key)  
 high = mid - 1;  
 **else  
 return** mid; *// key found* }  
 **return** -(low + 1); *// key not found.*}

5. Math6. 算法相关：PriorityQueue 基于堆实现的无界队列，非线程安全的An unbounded priority {@linkplain Queue queue} based on a priority heap.

Treemap基于红黑树实现

floorKey/floorEntry:返回稍小的元素

ceilingKey/ceilingEntry：返回稍大的元素

Guava有range版本。

# 经典数据结构/算法

## 排序

交换排序：冒泡、快排

插入、希尔

选择排序：简单选择、树形选择、堆排序

归并排序

基数排序

计数排序

桶排序

## 二叉树

### 性质

第i层有2^(i-1)个节点，i>=1.

一棵树至多有2^k-1个节点，k>=1。（引出满二叉树（满了，节点多），完全二叉树（最后一层不满，靠左）的定义）

完全二叉树的高度是log2n/1 +1

n0 = n2+1

数组表示法：从0开始

parent(i) = （i-1）/2 lchild(i) = 2i+1 rchild(i) = 2i+2

n个节点不相似二叉树1/(n+1)C(2n)(n)

### 构建

106. Construct Binary Tree from Inorder and Postorder Traversal

### 遍历

**public class** IterativeInOrder {  
 Stack<TreeNode> **stack** = **new** Stack<>();  
  
 **public void** run(TreeNode root){  
 *//初始化* pushAll(root);  
 **while**(!**stack**.empty()){  
 TreeNode visit = **stack**.pop();  
 doWork(visit);  
 pushAll(visit.**right**);  
 }  
 }  
  
 **public void** doWork(TreeNode visit){  
 System.***out***.println(visit.**val**);  
 }  
  
 */\*\*  
 \* 沿左树，一直压到叶子  
 \** ***@param root*** *\*/* **private void** pushAll(TreeNode root){  
 TreeNode thisNode = root;  
 **while**(thisNode!= **null**){  
 **stack**.push(thisNode);  
 thisNode = thisNode.**left**;  
 }  
 }  
}

### 二叉搜索树Binary Search Tree

二叉搜索树的线索可以利用空间加快搜索

后续：叶节点：右链；非叶节点：右子树坐下。

前驱：叶节点：左链；非叶节点：做子树右下。

习题：

https://leetcode.com/problems/unique-binary-search-trees/description/

C(2n,n)/(n+1)

<https://en.wikipedia.org/wiki/Catalan_number#Applications_in_combinatorics>

1. 完全二叉樹、二叉搜索树

Given n, how many structurally unique BST's (binary search trees) that store values 1 ... n?

unique-binary-search-trees-ii：如何生成具体的树：递归。左子树从0到n。（注意类似的情况都可以用catalan数计数。）

98. Validate Binary Search Tree：中根序遍历，递增

450. Delete Node in a BST

1. 叶子节点直接删除

2. 单孩节点：直接替代

3. 双孩节点：

方案一：前驱值替代，再删除前驱

方案二：左树替代，右树下降低到最低

230. Kth Smallest Element in a BST：中根序遍历，计数

What if the BST is modified (insert/delete operations) often and you need to find the kth smallest frequently? How would you optimize the kthSmallest routine?

增加一个计数域，同步更新

### 红黑树

理解要点

## 字典树

**public class** Node {  
 Node[] **children**;  
 String **leaf**;  
  
 */\*\*  
 \*  
 \** ***@param n*** *孩子容量  
 \*/* **public** Node(**int** n) {  
 **children** = **new** Node[n];  
 }  
}

**public class** Trie {  
 Node **root**;  
  
 **public** Node create(){  
 **root** = **new** Node(26);  
 **return root**;  
 }  
  
 **public void** insert(String s){  
 *//逐个查找，找不到则建点* Node node = **root**;  
 **for** (**int** i = 0; i < s.length(); i++) {  
 **int** index = order(s.charAt(0));  
 Node child = node.**children**[index];  
 **if**(child == **null**){  
 child = **new** Node(26);  
 node.**children**[index] = child;  
 node = child;  
 }**else**{  
 node = child;  
 }  
 }  
 node.**leaf** = s;  
 }  
  
 **public boolean** search(String s){  
 Node node = **root**;  
 **for** (**int** i = 0; i < s.length(); i++) {  
 **int** index = order(s.charAt(0));  
 Node child = node.**children**[index];  
 **if**(child == **null**){  
 **return false**;  
 }**else**{  
 node = child;  
 }  
 }  
 *//如果该词是其他词的前缀，也不存在。* **return** node.**leaf** != **null**;  
 }  
  
 **public int** order(**char** a){  
 **return** a - **'a'**;  
 }  
}

211. Add and Search Word - Data structure design

变种：search(word) can search a literal word or a regular expression string containing only letters a-z or .. A . means it can represent any one letter.

648. Replace Words

676. Implement Magic Dictionary

677. Map Sum Pairs

红黑树

前缀和

线段树

图

拓扑排序

并查集扩展元素数量集合数量

排序

计算几何

算法导论33章

字符串匹配

kmp算法ac自动机

## 素数计算

650. 2 Keys Keyboard

Initially on a notepad only one character 'A' is present. You can perform two operations on this notepad for each step:

1. Copy All: You can copy all the characters present on the notepad (partial copy is not allowed).
2. Paste: You can paste the characters which are copied **last time**.

Given a number n. You have to get **exactly** n 'A' on the notepad by performing the minimum number of steps permitted. Output the minimum number of steps to get n 'A'.

解：n = (CP…)(CPP…)(CPP…)=>分解n的因子

考虑p+q <= pq=》求质因子和

## 二分查找

含重复元素、等号、上下界

34. Search for a Range

## 流算法

### 摩尔多数投票算法

查找1/2水王

The algorithm maintains in its [local variables](https://en.wikipedia.org/wiki/Local_variable) a sequence element and a counter, with the counter initially zero. It then processes the elements of the sequence, one at a time. When processing an element *x*, if the counter is zero, the algorithm stores *x* as its remembered sequence element and sets the counter to one. Otherwise, it compares *x* to the stored element and either increments the counter (if they are equal) or decrements the counter (otherwise). At the end of this process, if the sequence has a majority, it will be the element stored by the algorithm. This can be expressed in [pseudocode](https://en.wikipedia.org/wiki/Pseudocode) as the following steps:

* Initialize an element *m* and a counter *i* with *i* = 0
* For each element *x* of the input sequence:
  + If *i* = 0, then assign *m* = *x* and *i* = 1
  + else if *m* = *x*, then assign *i* = *i* + 1
  + else assign *i* = *i* − 1
* Return *m*

Even when the input sequence has no majority, the algorithm will report one of the sequence elements as its result. However, it is possible to perform a second pass over the same input sequence in order to count the number of times the reported element occurs and determine whether it is actually a majority. This second pass is needed, as it is not possible for a sublinear-space algorithm to determine whether there exists a majority element in a single pass through the input

另一种思路：配对删除。最终状态：留下一个（就是水王）；两个（必然两个水王）

1/3水王

The basic idea is based on Moore's Voting Algorithm, we need two candidates with top 2 frequency. If meeting different number from the candidate, then decrease 1 from its count, or increase 1 on the opposite condition. Once count equals 0, then switch the candidate to the current number. The trick is that we need to count again for the two candidates after the first loop. Finally, output the numbers appearing more than n/3 times.

另一种思路：配对删除。最终状态：如果最终留下两个元素，就无法判断了，所以还需要再遍历一边。

### 水库采样

目的在于从包含n个项目的集合S中选取k个样本，其中n为一很大或未知的数量，尤其适用于不能把所有n个项目都存放到主内存的情况。

在高德纳的计算机程序设计艺术中，有如下问题：**可否在一未知大小的集合中，随机取出一元素？**。或者是Google面试题： I have a linked list of numbers of length N. N is very large and I don’t know in advance the exact value of N. How can I most efficiently write a function that will return k completely random numbers from the list（中文简化的意思就是：**在不知道文件总行数的情况下，如何从文件中随机的抽取一行？**）。两题的核心意思都是在总数不知道的情况下如何等概率地从中抽取一行？即是说如果最后发现文字档共有N行，则每一行被抽取的概率均为1/N？

我们可以：定义取出的行号为choice，第一次直接以第一行作为取出行 choice ，而后第二次以二分之一概率决定是否用第二行替换 choice ，第三次以三分之一的概率决定是否以第三行替换 choice ……，以此类推。由上面的分析我们可以得出结论，**在取第n个数据的时候，我们生成一个0到1的随机数p，如果p小于1/n，保留第n个数。大于1/n，继续保留前面的数。直到数据流结束，返回此数，算法结束。**

**问题一**

首先考虑k为1的情况，即：给定一个长度很大或者长度未知数据流，限定对每个元素只能访问一次，写出一个随机选择算法，使得所有元素被选中的概率相等。

设当前读取的是第n个元素，采用归纳法分析如下：

1. n = 1 时，只有一个元素，直接返回即可，概率为1。
2. n = 2 时，需要等概率返回前两个元素，显然概率为1/2。可以生成一个0～1之间的随机数p，p < 0.5 时返回第一个，否则返回第二个。
3. n = 3 时，要求每个元素返回的概率为1/3。注意此时前两个元素留下来的概率均为1/2。做法是：生成一个0～1之间的随机数，若<1/3，则返回第三个，否则返回上一步留下的那个。元素1和2留下的概率均为：1/2 \* (1 - 1/3) = 1/3，即上一步留下的概率乘以这一步留下（即元素3不留下）的概率。
4. 假设 n = m 时，前n个元素留下的概率均为：1/n = 1/m；
5. 那么 n = m+1 时，生成0～1之间的随机数并判断是否<1/(m+1)，若是则留下元素m+1，否则留下上一步留下的元素。这样一来，元素m+1留下的概率为1/(m+1)，前m个元素留下来的概率均为：1/m \* (1 - 1/(m+1)) = 1/(m+1)，也就是1/n。
6. 综上可知，算法成立。

**问题二**

将问题一中的条件变为，k为任意整数的情况，即要求最终返回的元素有k个，这就是水塘抽样（Reservoir Sampling）问题。要求是：取到第n个元素时，前n个元素被留下的几率相等，即k/n。

算法类似，将1/n换乘**k/n**即可。在取第n个数据的时候，我们生成一个0到1的随机数p，如果p小于k/n，替换池中任意一个为第n个数。大于k/n，继续保留前面的数。直到数据流结束，返回此k个数。但是**为了保证计算机计算分数额准确性，一般是生成一个0到n的随机数，跟k相比，道理是一样的**。

同样采用归纳法来分析：

1. 初始情况 n <= k：此时每个元素留下的概率均为1。
2. 当 n = k+1 时，第k+1个元素留下的概率为k/(k+1)，前k个元素留下的概率均为：k/k \* (1 - k/(k+1) \* 1/k) = k/(k+1)，即上一步留下的概率乘以这一步留下的概率。
3. 假设 n = m 时，每个元素留下的概率均为 k/n = k/m。
4. 那么，当 n = m+1 时，第m+1个元素留下的概率为1/(m+1)，前m个元素留下的概率均为：k/m \* (1 - k/(m+1) \* 1/k) = k/(m+1)，其中：k/m为上一步留下来的概率，k/(m+1) \* 1/k 为这一步不能留下来的概率（第m+1个留下来，同时池中一个元素被踢出的概率）。
5. 综上可知，算法成立。

伪代码如下：

//stream代表数据流

//reservoir代表返回长度为k的池塘

//从stream中取前k个放入reservoir；

for ( int i = 1; i < k; i++)

reservoir[i] = stream[i];

for (i = k; stream != null; i++) {

p = random(0, i);

if (p < k) reservoir[p] = stream[i];

return reservoir;

## 下一个排列

O(n);O(1)

**public** **void** **nextPermutation(int[]** nums**)** **{**

**int** i **=** nums**.**length **-** 2**;**

**while** **(**i **>=** 0 **&&** nums**[**i **+** 1**]** **<=** nums**[**i**])** **{ //找到第一个不满足逆序的数**

i**--;**

**}**

**if** **(**i **>=** 0**)** **{**

**int** j **=** nums**.**length **-** 1**;**

**while** **(**j **>=** 0 **&&** nums**[**j**]** **<=** nums**[**i**])** **{ //找到比i位置大一点的数**

j**--;**

**}**

swap**(**nums**,** i**,** j**); //i位置确定好**

**}**

reverse**(**nums**,** i **+** 1**); //余下的按正序排列即可**

**}**

## 树状数组/binary indexed tree/ Fenwick tree

https://blog.csdn.net/l664675249/article/details/50157669

**问题：求一个数组中连续n项的和。**

首先想到的肯定是做一个循环，把这个连续的n项加起来，时间复杂度为O（n）。

**会不会有O（logn）的解法？**

如果只操作一次还是可以接受的，但是如果需要大量的求和操作，比如第一次求下标（1，1234）的和第二次求下标（2，1024）的和，很容易发现在第一次计算的过程中（2，1024）的和是计算过的，只是没有保存下来，导致第二次求和的时候还要再算一遍。如果事先把一部分的和先计算并保存起来，这样会不会更快一些呢？

树状数组是一个查询和修改复杂度都为log(n)的数据结构。主要用于查询任意两位之间的所有元素之和，但是每次只能修改一个元素的值。

**核心思想:**

* 树状数组中的每个元素是原数组中一个或者多个连续元素的和。
* 在进行连续求和操作a[1]+…+a[n]时，只需要将树状数组中某几个元素的和即可。时间复杂度为O(lgn)

下面是一个示意图



a[]: 保存原始数据的数组   
e[]: 树状数组，其中的任意一个元素e[i]可能是一个或者多个a数组中元素的和。如e[2]=a[1]+a[2]; e[3]=a[3]，e[4]=a[1]+a[2]+a[3]+a[4]。   
e[i]中的元素：如果数字 i 的二进制表示中末尾有**k个连续的0，则e[i]是a数组中2^k个元素的和**，则e[i]=a[**i-2^k+1**]+a[i-2^k+2]+…+a[i-1]+a[i]。也就是说，**e[i]中每一个元素管理着a[]中若干个元素的和，并且各个元素管理的区间没有重叠。**

　　　　如：4=100(2)　　e[4]=a[1]+a[2]+a[3]+a[4];   
　　　　　　6=110(2)　　e[6]=a[5]+a[6]   
　　　　　　7=111(2)　　e[7]=a[7]   
　　　　　　   
计算2^k的两个方法

* 2^k = (i & (-i)); (利用机器补码特性)
* 2^k = (i & (i^(i-1));

**父节点**

是离它最近的，且编号末位连续0比它多的就是它的父亲,如e[2]是e[1]的儿子；e[4]是e[2]的儿子。   
e[4] = e[2]+e[3]+a[4] = a[1]+a[2]+a[3]+a[4] ，e[2]、e[3]的后继就是e[4]。

**计算方法**

lowbit(i) = ( (i-1) ^ i) & i ; //或者(i & (-i))   
**节点e[i]的父节点为 e[ i - lowbit(i) ]**

**子节点**

最近的，编号即为比自己小的，最末连续0比自己多的节点。如e[7]的子节点是e[6],e[6]的子节点是e[4]

**计算方法**

lowbit(i) = ( (i-1) ^ i) & i ; //或者(i & (-i))   
**节点e[i]的子节点为 e[ i + lowbit(i) ]**

**实现代码**

public class NumArray {

private int[] tree; //Binary Indexed Tree

private int[] nums; //原始数组

public NumArray(int[] nums) {

this.nums = nums;

int sum = 0;

int lowbit;

tree = new int[nums.length + 1];

for (int i = 1; i < tree.length; i++) {

sum = 0;

lowbit = i & ((i - 1) ^ i);

for (int j = i; j > i - lowbit; j--) {

sum = sum + nums[j - 1];

}

tree[i] = sum;

}

}

//更新

void update(int i, int val) {

int tem = val - nums[i];

nums[i] = val;

i++;

for (; i < tree.length; i = i + (i & ((i - 1) ^ i))) {

tree[i] += tem;

}

}

public int sumRange(int i, int j) {

return getSum(j) - getSum(i - 1);

}

//求和

public int getSum(int i) {

int sum = 0;

i++;

while (i > 0) {

sum = sum + tree[i];

i = i - (i & ((i - 1) ^ i));

}

return sum;

}

}

## 树

树的表示法

1双亲表示法：数组

2 孩子表示法

3 孩子兄弟表示法

310. Minimum Height Trees

a tree is an undirected graph in which any two vertices are connected by exactly one path. In other words, any connected graph without simple cycles is a tree.

//确定树根

We start from every end, by end we mean vertex of degree 1 (aka leaves). We let the pointers move the same speed. When two pointers meet, we keep only one of them, until the last two pointers meet or one step away we then find the roots.

## 线段树

307. Range Sum Query – Mutable

<https://leetcode.com/problems/range-sum-query-mutable/solution/>

https://en.wikipedia.org/wiki/Segment\_tree

Segment tree is a very flexible data structure, because it is used to solve numerous range query problems like finding minimum, maximum, sum, greatest common divisor, least common denominator in array in logarithmic time.

这图是不是有问题？

The segment tree for array *a*[0,1,…,*n*−1] is a binary tree in which each node contains **aggregate**information (min, max, sum, etc.) for a subrange [*i*…*j*] of the array, as its left and right child hold information for range [i \ldots \frac{i+j}{2}][*i*…​2​​*i*+*j*​​] and [\frac{i + j}{2} + 1, j][​2​​*i*+*j*​​+1,*j*].

Segment tree could be implemented using either an array or a tree. For an array implementation, if the element at index iis not a leaf, its left and right child are stored at index 2iand 2i + 1respectively.

In the example above (Figure 2), every leaf node contains the initial array elements {2,4,5,7,8,9}. The internal nodes contain the sum of the corresponding elements in range - (11) for the elements from index 0 to index 2. The root (35) being the sum of its children (6);(29), holds the total sum of the entire array.

Segment Tree can be broken down to the three following steps:

1. Pre-processing step which builds the segment tree from a given array.
2. Update the segment tree when an element is modified.
3. Calculate the Range Sum Query using the segment tree.

1. Build segment tree：bottom-up approach

We already know from the above that if some node p*p* holds the sum of [*i*…*j*] range, its left and right children hold the sum for range [i \ldots \frac{i + j}{2}][*i*…​2​​*i*+*j*​​] and [\frac{i + j}{2} + 1, j][​2​​*i*+*j*​​+1,*j*] respectively.

Therefore to find the sum of node p, we need to calculate the sum of its right and left child in advance.

We begin from the leaves, initialize them with input array elements *a*[0,1,…,*n*−1]. Then we move upward to the higher level to calculate the parents' sum till we get to the root of the segment tree.

//注意从1开始

**int[]** tree**;**

**int** n**;**

**public** **NumArray(int[]** nums**)** **{**

n **=** nums**.**length**;**

tree **=** **new** **int[**n **\*** 2**];**

**for** **(int** i **=** n**,** j **=** 0**;** i **<** 2 **\*** n**;** i**++,** j**++)**

tree**[**i**]** **=** nums**[**j**];//后面的一半是叶子节点**

**for** **(int** i **=** n **-** 1**;** i **>** 0**;** **--**i**)**

tree**[**i**]** **=** tree**[**i **\*** 2**]** **+** tree**[**i **\*** 2 **+** 1**];//主干节点**

**}**

**Complexity Analysis**

* Time complexity :*O*(*n*)

because we calculate the sum of one node during each iteration of the for loop. There are approximately2*n* nodes.

This could be proved in the following way: Segmented tree for array with nelements has n*n* leaves (the array elements itself). The number of nodes in each level is half the number in the level below.

So if we sum the number by level we will get:

n + n/2 + n/4 + n/8 + \ldots + 1 \approx 2n*n*+*n*/2+*n*/4+*n*/8+…+1≈2*n*

* Space complexity : O(n).

2. Update segment tree

When we update the array at some index i*i* we need to rebuild the segment tree, because there are tree nodes which contain the sum of the modified element. Again we will use a bottom-up approach. We update the leaf node that stores a[i]. From there we will follow the path up to the root updating the value of each parent as a sum of its children values.

**void** **update(int** pos**,** **int** val**)** **{**

pos **+=** n**;**

tree**[**pos**]** **=** val**;**

**while** **(**pos **>** 0**)** **{**

**int** left **=** pos**;**

**int** right **=** pos**;**

**if** **(**pos **%** 2 **==** 0**)** **{**

right **=** pos **+** 1**;**

**}** **else** **{**

left **=** pos **-** 1**;**

**}**

*// parent is updated after child is updated*

tree**[**pos **/** 2**]** **=** tree**[**left**]** **+** tree**[**right**];**

pos **/=** 2**;**

**}**

**}**

**Complexity Analysis**

* Time complexity : *O*(log*n*).

because there are a few tree nodes with range that include ith array element, one on each level. There are log(*n*) levels.

* Space complexity : O(1)*O*(1).

3. Range Sum Query

We can find range sum query [L, R]:

Algorithm hold loop invariant:

l \le r*l*≤*r* and sum of [*L*…*l*] and [*r*…*R*] has been calculated, where land rare the left and right boundary of calculated sum. Initially we set lwith left leaf L*L* and r*r* with right leaf R*R*. Range [l, r] hrinks on each iteration till range borders meets after approximately \log nlog*n* iterations of the algorithm

* Loop till l \le r*l*≤*r*
  + Check if l*l* is right child of its parent P*P*
    - l*l* is right child of P*P*. Then P*P* contains sum of range of l*l* and another child which is outside the range [l, r][*l*,*r*] and we don't need parent P*P* sum. Add l*l* to sum*sum* without its parent P*P* and set l*l* to point to the right of P*P* on the upper level.
    - l*l* is not right child of P*P*. Then parent P*P* contains sum of range which lies in [l, r][*l*,*r*]. Add P*P* to sum*sum*and set l*l* to point to the parent of P*P*
  + Check if r*r* is left child of its parent P*P*
    - r*r* is left child of P*P*. Then P*P* contains sum of range of r*r* and another child which is outside the range [l, r][*l*,*r*] and we don't need parent P*P* sum. Add r*r* to sum*sum* without its parent P*P* and set r*r* to point to the left of P*P* on the upper level.
    - r*r* is not left child of P*P*. Then parent P*P* contains sum of range which lies in [l, r][*l*,*r*]. Add P*P* to sum*sum*and set r*r* to point to the parent of P*P*

**public** **int** **sumRange(int** l**,** **int** r**)** **{**

*// get leaf with value 'l'*

l **+=** n**;**

*// get leaf with value 'r'*

r **+=** n**;**

**int** sum **=** 0**;**

**while** **(**l **<=** r**)** **{**

**if** **((**l **%** 2**)** **==** 1**)** **{**

sum **+=** tree**[**l**];**

l**++;**

**}**

**if** **((**r **%** 2**)** **==** 0**)** **{**

sum **+=** tree**[**r**];**

r**--;**

**}**

**l /= 2;**

**r /= 2;**

**}**

**return** sum**;**

**}**

**Complexity Analysis**

* Time complexity : O(\log n)*O*(log*n*)

Time complexity is O(\log n)*O*(log*n*) because on each iteration of the algorithm we move one level up, either to the parent of the current node or to the next sibling of parent to the left or right direction till the two boundaries meet. In the worst-case scenario this happens at the root after \log nlog*n* iterations of the algorithm.

* Space complexity : O(1)*O*(1).

## 分桶法和平方分割

<https://leetcode.com/problems/range-sum-query-mutable/solution/>

思路就是避免重复计算

分割成sqrt（n）个桶

## 图

### 图的表示

1 2D数组

2 邻接表：顶点数组+边链表

3 十字链表：邻接表+逆邻接表

4 邻接多重表

133. Clone Graph

节点映射，记录对应关系？

### 二分图

Bipartite graph

From Wikipedia, the free encyclopedia

[](https://en.wikipedia.org/wiki/File:Simple-bipartite-graph.svg)

Example of a bipartite graph without cycles

[](https://en.wikipedia.org/wiki/File:Biclique_K_3_5.svg)

A [complete bipartite graph](https://en.wikipedia.org/wiki/Complete_bipartite_graph) with m = 5 and n = 3

In the [mathematical](https://en.wikipedia.org/wiki/Mathematics) field of [graph theory](https://en.wikipedia.org/wiki/Graph_theory), a **bipartite graph** (or **bigraph**) is a [graph](https://en.wikipedia.org/wiki/Graph_(discrete_mathematics)) whose [vertices](https://en.wikipedia.org/wiki/Vertex_(graph_theory)) can be divided into two [disjoint](https://en.wikipedia.org/wiki/Disjoint_sets) and [independent sets](https://en.wikipedia.org/wiki/Independent_set_(graph_theory)) {\displaystyle U} and {\displaystyle V}such that every [edge](https://en.wikipedia.org/wiki/Edge_(graph_theory)) connects a vertex in {\displaystyle U} to one in {\displaystyle V}. Vertex sets {\displaystyle U} and {\displaystyle V} are usually called the *parts* of the graph. Equivalently, a bipartite graph is a graph that does not contain any odd-length [cycles](https://en.wikipedia.org/wiki/Cycle_(graph_theory)).[[1]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-diestel2005graph-1)[[2]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-2)

The two sets {\displaystyle U} and {\displaystyle V} may be thought of as a [coloring](https://en.wikipedia.org/wiki/Graph_coloring) of the graph with two colors: if one colors all nodes in {\displaystyle U} blue, and all nodes in {\displaystyle V} green, each edge has endpoints of differing colors, as is required in the graph coloring problem.[[3]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-adh98-7-3)[[4]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-s12-4) In contrast, such a coloring is impossible in the case of a non-bipartite graph, such as a [triangle](https://en.wikipedia.org/wiki/Gallery_of_named_graphs): after one node is colored blue and another green, the third vertex of the triangle is connected to vertices of both colors, preventing it from being assigned either color.

One often writes {\displaystyle G=(U,V,E)} to denote a bipartite graph whose partition has the parts {\displaystyle U} and {\displaystyle V}, with {\displaystyle E} denoting the edges of the graph. If a bipartite graph is not [connected](https://en.wikipedia.org/wiki/Connected_graph), it may have more than one bipartition;[[5]](https://en.wikipedia.org/wiki/Bipartite_graph" \l "cite_note-5) in this case, the {\displaystyle (U,V,E)} notation is helpful in specifying one particular bipartition that may be of importance in an application. If {\displaystyle |U|=|V|}, that is, if the two subsets have equal [cardinality](https://en.wikipedia.org/wiki/Cardinality), then {\displaystyle G} is called a *balanced* bipartite graph.[[3]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-adh98-7-3) If all vertices on the same side of the bipartition have the same [degree](https://en.wikipedia.org/wiki/Degree_(graph_theory)), then {\displaystyle G} is called [biregular](https://en.wikipedia.org/wiki/Biregular_graph).

**Contents**

  [hide]

* [1Examples](https://en.wikipedia.org/wiki/Bipartite_graph#Examples)
* [2Properties](https://en.wikipedia.org/wiki/Bipartite_graph#Properties)
  + [2.1Characterization](https://en.wikipedia.org/wiki/Bipartite_graph#Characterization)
  + [2.2König's theorem and perfect graphs](https://en.wikipedia.org/wiki/Bipartite_graph#K%C3%B6nig's_theorem_and_perfect_graphs)
  + [2.3Degree](https://en.wikipedia.org/wiki/Bipartite_graph#Degree)
  + [2.4Relation to hypergraphs and directed graphs](https://en.wikipedia.org/wiki/Bipartite_graph#Relation_to_hypergraphs_and_directed_graphs)
* [3Algorithms](https://en.wikipedia.org/wiki/Bipartite_graph#Algorithms)
  + [3.1Testing bipartiteness](https://en.wikipedia.org/wiki/Bipartite_graph#Testing_bipartiteness)
  + [3.2Odd cycle transversal](https://en.wikipedia.org/wiki/Bipartite_graph#Odd_cycle_transversal)
  + [3.3Matching](https://en.wikipedia.org/wiki/Bipartite_graph#Matching)
* [4Additional applications](https://en.wikipedia.org/wiki/Bipartite_graph#Additional_applications)
* [5See also](https://en.wikipedia.org/wiki/Bipartite_graph#See_also)
* [6References](https://en.wikipedia.org/wiki/Bipartite_graph#References)
* [7External links](https://en.wikipedia.org/wiki/Bipartite_graph#External_links)

Examples[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=1" \o "Edit section: Examples)]

When modelling relations between two different classes of objects, bipartite graphs very often arise naturally. For instance, a graph of football players and clubs, with an edge between a player and a club if the player has played for that club, is a natural example of an *affiliation network*, a type of bipartite graph used in [social network analysis](https://en.wikipedia.org/wiki/Social_network_analysis).[[6]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-6)

Another example where bipartite graphs appear naturally is in the ([NP-complete](https://en.wikipedia.org/wiki/NP-complete)) railway optimization problem, in which the input is a schedule of trains and their stops, and the goal is to find a set of train stations as small as possible such that every train visits at least one of the chosen stations. This problem can be modeled as a [dominating set](https://en.wikipedia.org/wiki/Dominating_set) problem in a bipartite graph that has a vertex for each train and each station and an edge for each pair of a station and a train that stops at that station.[[7]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-niedermeier2006invitiation-7)

A third example is in the academic field of numismatics. Ancient coins are made using two positive impressions of the design (the obverse and reverse). The charts numismatists produce to represent the production of coins are bipartite graphs. [[8]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-bracey2012-8)

More abstract examples include the following:

* Every [tree](https://en.wikipedia.org/wiki/Tree_(graph_theory)) is bipartite.[[4]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-s12-4)
* [Cycle graphs](https://en.wikipedia.org/wiki/Cycle_graph) with an even number of vertices are bipartite.[[4]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-s12-4)
* Every [planar graph](https://en.wikipedia.org/wiki/Planar_graph) whose [faces](https://en.wikipedia.org/wiki/Glossary_of_graph_theory#Genus) all have even length is bipartite.[[9]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-9) Special cases of this are [grid graphs](https://en.wikipedia.org/wiki/Grid_graph) and [squaregraphs](https://en.wikipedia.org/wiki/Squaregraph), in which every inner face consists of 4 edges and every inner vertex has four or more neighbors.[[10]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-10)
* The [complete bipartite graph](https://en.wikipedia.org/wiki/Complete_bipartite_graph) on *m* and *n* vertices, denoted by *Kn,m* is the bipartite graph {\displaystyle G=(U,V,E)}, where *U* and *V* are disjoint sets of size *m* and *n*, respectively, and *E* connects every vertex in *U* with all vertices in *V*. It follows that *Km,n* has *mn* edges.[[11]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-11) Closely related to the complete bipartite graphs are the [crown graphs](https://en.wikipedia.org/wiki/Crown_graph), formed from complete bipartite graphs by removing the edges of a [perfect matching](https://en.wikipedia.org/wiki/Perfect_matching).[[12]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-12)
* [Hypercube graphs](https://en.wikipedia.org/wiki/Hypercube_graph), [partial cubes](https://en.wikipedia.org/wiki/Partial_cube), and [median graphs](https://en.wikipedia.org/wiki/Median_graph) are bipartite. In these graphs, the vertices may be labeled by [bitvectors](https://en.wikipedia.org/wiki/Bitvector), in such a way that two vertices are adjacent if and only if the corresponding bitvectors differ in a single position. A bipartition may be formed by separating the vertices whose bitvectors have an even number of ones from the vertices with an odd number of ones. Trees and squaregraphs form examples of median graphs, and every median graph is a partial cube.[[13]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-13)

Properties[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=2" \o "Edit section: Properties)]

**Characterization**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=3" \o "Edit section: Characterization)]

Bipartite graphs may be characterized in several different ways:

* A graph is bipartite [if and only if](https://en.wikipedia.org/wiki/If_and_only_if) it does not contain an [odd cycle](https://en.wikipedia.org/wiki/Cycle_(graph_theory)).[[14]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-14)
* A graph is bipartite if and only if it is 2-colorable, (i.e. its [chromatic number](https://en.wikipedia.org/wiki/Chromatic_number) is less than or equal to 2).[[3]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-adh98-7-3)
* The [spectrum](https://en.wikipedia.org/wiki/Spectral_graph_theory) of a graph is symmetric if and only if it's a bipartite graph.[[15]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-15)

**König's theorem and perfect graphs**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=4" \o "Edit section: König's theorem and perfect graphs)]

In bipartite graphs, the size of [minimum vertex cover](https://en.wikipedia.org/wiki/Minimum_vertex_cover) is equal to the size of the [maximum matching](https://en.wikipedia.org/wiki/Maximum_matching); this is [König's theorem](https://en.wikipedia.org/wiki/K%C3%B6nig%27s_theorem_(graph_theory)).[[16]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-16)[[17]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-17) An alternative and equivalent form of this theorem is that the size of the [maximum independent set](https://en.wikipedia.org/wiki/Maximum_independent_set) plus the size of the maximum matching is equal to the number of vertices. In any graph without [isolated vertices](https://en.wikipedia.org/wiki/Isolated_vertex) the size of the [minimum edge cover](https://en.wikipedia.org/wiki/Minimum_edge_cover) plus the size of a maximum matching equals the number of vertices.[[18]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-18) Combining this equality with König's theorem leads to the facts that, in bipartite graphs, the size of the minimum edge cover is equal to the size of the maximum independent set, and the size of the minimum edge cover plus the size of the minimum vertex cover is equal to the number of vertices.

Another class of related results concerns [perfect graphs](https://en.wikipedia.org/wiki/Perfect_graph): every bipartite graph, the [complement](https://en.wikipedia.org/wiki/Complement_(graph_theory)) of every bipartite graph, the [line graph](https://en.wikipedia.org/wiki/Line_graph) of every bipartite graph, and the complement of the line graph of every bipartite graph, are all perfect. Perfection of bipartite graphs is easy to see (their [chromatic number](https://en.wikipedia.org/wiki/Chromatic_number) is two and their [maximum clique](https://en.wikipedia.org/wiki/Maximum_clique) size is also two) but perfection of the [complements](https://en.wikipedia.org/wiki/Complement_(graph_theory)) of bipartite graphs is less trivial, and is another restatement of König's theorem. This was one of the results that motivated the initial definition of perfect graphs.[[19]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-19) Perfection of the complements of line graphs of perfect graphs is yet another restatement of König's theorem, and perfection of the line graphs themselves is a restatement of an earlier theorem of König, that every bipartite graph has an [edge coloring](https://en.wikipedia.org/wiki/Edge_coloring) using a number of colors equal to its maximum degree.

According to the [strong perfect graph theorem](https://en.wikipedia.org/wiki/Strong_perfect_graph_theorem), the perfect graphs have a [forbidden graph characterization](https://en.wikipedia.org/wiki/Forbidden_graph_characterization) resembling that of bipartite graphs: a graph is bipartite if and only if it has no odd cycle as a subgraph, and a graph is perfect if and only if it has no odd cycle or its [complement](https://en.wikipedia.org/wiki/Complement_(graph_theory)) as an [induced subgraph](https://en.wikipedia.org/wiki/Induced_subgraph). The bipartite graphs, line graphs of bipartite graphs, and their complements form four out of the five basic classes of perfect graphs used in the proof of the strong perfect graph theorem.[[20]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-20)

**Degree**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=5" \o "Edit section: Degree)]

For a vertex, the number of adjacent vertices is called the [degree](https://en.wikipedia.org/wiki/Degree_(graph_theory)) of the vertex and is denoted {\displaystyle \deg(v)}. The *degree sum formula* for a bipartite graph states that

{\displaystyle \sum \_{v\in V}\deg(v)=\sum \_{u\in U}\deg(u)=|E|\,.}

The degree sequence of a bipartite graph is the pair of lists each containing the degrees of the two parts {\displaystyle U} and {\displaystyle V}. For example, the complete bipartite graph *K*3,5 has degree sequence {\displaystyle (5,5,5),(3,3,3,3,3)}. Isomorphic bipartite graphs have the same degree sequence. However, the degree sequence does not, in general, uniquely identify a bipartite graph; in some cases, non-isomorphic bipartite graphs may have the same degree sequence.

The [bipartite realization problem](https://en.wikipedia.org/wiki/Bipartite_realization_problem) is the problem of finding a simple bipartite graph with the degree sequence being two given lists of natural numbers. (Trailing zeros may be ignored since they are trivially realized by adding an appropriate number of isolated vertices to the digraph.)

**Relation to hypergraphs and directed graphs**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=6" \o "Edit section: Relation to hypergraphs and directed graphs)]

The [biadjacency matrix](https://en.wikipedia.org/wiki/Adjacency_matrix_of_a_bipartite_graph) of a bipartite graph {\displaystyle (U,V,E)} is a [(0,1) matrix](https://en.wikipedia.org/wiki/(0,1)_matrix) of size {\displaystyle |U|\times |V|} that has a one for each pair of adjacent vertices and a zero for nonadjacent vertices.[[21]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-21) Biadjacency matrices may be used to describe equivalences between bipartite graphs, hypergraphs, and directed graphs.

A [hypergraph](https://en.wikipedia.org/wiki/Hypergraph) is a combinatorial structure that, like an undirected graph, has vertices and edges, but in which the edges may be arbitrary sets of vertices rather than having to have exactly two endpoints. A bipartite graph {\displaystyle (U,V,E)} may be used to model a hypergraph in which *U* is the set of vertices of the hypergraph, *V* is the set of hyperedges, and *E* contains an edge from a hypergraph vertex *v* to a hypergraph edge *e* exactly when *v* is one of the endpoints of *e*. Under this correspondence, the biadjacency matrices of bipartite graphs are exactly the [incidence matrices](https://en.wikipedia.org/wiki/Incidence_matrix) of the corresponding hypergraphs. As a special case of this correspondence between bipartite graphs and hypergraphs, any [multigraph](https://en.wikipedia.org/wiki/Multigraph) (a graph in which there may be two or more edges between the same two vertices) may be interpreted as a hypergraph in which some hyperedges have equal sets of endpoints, and represented by a bipartite graph that does not have multiple adjacencies and in which the vertices on one side of the bipartition all have [degree](https://en.wikipedia.org/wiki/Degree_(graph_theory)) two.[[22]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-22)

A similar reinterpretation of adjacency matrices may be used to show a one-to-one correspondence between [directed graphs](https://en.wikipedia.org/wiki/Directed_graph) (on a given number of labeled vertices, allowing self-loops) and balanced bipartite graphs, with the same number of vertices on both sides of the bipartition. For, the adjacency matrix of a directed graph with *n* vertices can be any [(0,1) matrix](https://en.wikipedia.org/wiki/(0,1)_matrix) of size {\displaystyle n\times n}, which can then be reinterpreted as the adjacency matrix of a bipartite graph with *n* vertices on each side of its bipartition.[[23]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-23) In this construction, the bipartite graph is the [bipartite double cover](https://en.wikipedia.org/wiki/Bipartite_double_cover) of the directed graph.

Algorithms[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=7" \o "Edit section: Algorithms)]

**Testing bipartiteness**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=8" \o "Edit section: Testing bipartiteness)]

It is possible to test whether a graph is bipartite, and to return either a two-coloring (if it is bipartite) or an odd cycle (if it is not) in [linear time](https://en.wikipedia.org/wiki/Linear_time), using [depth-first search](https://en.wikipedia.org/wiki/Depth-first_search).

1. The main idea is to assign to each vertex the color that differs from the color of its parent in the depth-first search forest, assigning colors in a [preorder traversal](https://en.wikipedia.org/wiki/Preorder_traversal) of the depth-first-search forest. This will necessarily provide a two-coloring of the [spanning forest](https://en.wikipedia.org/wiki/Spanning_forest) consisting of the edges connecting vertices to their parents, but it may not properly color some of the non-forest edges. In a depth-first search forest, one of the two endpoints of every non-forest edge is an ancestor of the other endpoint, and when the depth first search discovers an edge of this type it should check that these two vertices have different colors. If they do not, then the path in the forest from ancestor to descendant, together with the miscolored edge, form an odd cycle, which is returned from the algorithm together with the result that the graph is not bipartite. However, if the algorithm terminates without detecting an odd cycle of this type, then every edge must be properly colored, and the algorithm returns the coloring together with the result that the graph is bipartite.[[24]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-24)

2.  [breadth-first search](https://en.wikipedia.org/wiki/Breadth-first_search) . Again, each node is given the opposite color to its parent in the search forest. If, when a vertex is colored, there exists an edge connecting it to a previously-colored vertex with the same color, then this edge together with the paths in the breadth-first search forest connecting its two endpoints to their [lowest common ancestor](https://en.wikipedia.org/wiki/Lowest_common_ancestor) forms an odd cycle. If the algorithm terminates without finding an odd cycle in this way, then it must have found a proper coloring, and can safely conclude that the graph is bipartite.[[25]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-25)

For the [intersection graphs](https://en.wikipedia.org/wiki/Intersection_graph) of {\displaystyle n} [line segments](https://en.wikipedia.org/wiki/Line_segment) or other simple shapes in the [Euclidean plane](https://en.wikipedia.org/wiki/Euclidean_plane), it is possible to test whether the graph is bipartite and return either a two-coloring or an odd cycle in time {\displaystyle O(n\log n)}, even though the graph itself may have upto {\displaystyle O(n^{2})} edges.[[26]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-26)

**Odd cycle transversal**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=9" \o "Edit section: Odd cycle transversal)]

*Main article:*[*Odd cycle transversal*](https://en.wikipedia.org/wiki/Odd_cycle_transversal)

[](https://en.wikipedia.org/wiki/File:Odd_Cycle_Transversal_of_size_2.png)

A graph with an odd cycle transversal of size 2: removing the two blue bottom vertices leaves a bipartite graph.

[Odd cycle transversal](https://en.wikipedia.org/wiki/Odd_cycle_transversal) is an [NP-complete](https://en.wikipedia.org/wiki/NP-complete) [algorithmic](https://en.wikipedia.org/wiki/Algorithm) problem that asks, given a graph *G* = (*V*,*E*) and a number *k*, whether there exists a set of *k* vertices whose removal from *G* would cause the resulting graph to be bipartite.[[27]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-yannakakis1978node-27) The problem is [fixed-parameter tractable](https://en.wikipedia.org/wiki/Parameterized_complexity), meaning that there is an algorithm whose running time can be bounded by a polynomial function of the size of the graph multiplied by a larger function of *k*.[[28]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-reed2004finding-28) The name *odd cycle transversal* comes from the fact that a graph is bipartite if and only if it has no odd [cycles](https://en.wikipedia.org/wiki/Cycle_(graph_theory)). Hence, to delete vertices from a graph in order to obtain a bipartite graph, one needs to "hit all odd cycle", or find a so-called odd cycle [transversal](https://en.wikipedia.org/wiki/Transversal_(combinatorics)) set. In the illustration, every odd cycle in the graph contains the blue (the bottommost) vertices, so removing those vertices kills all odd cycles and leaves a bipartite graph.

The *edge bipartization* problem is the algorithmic problem of deleting as few edges as possible to make a graph bipartite and is also an important problem in graph modification algorithmics. This problem is also [fixed-parameter tractable](https://en.wikipedia.org/wiki/Fixed-parameter_tractable), and can be solved in time *O*(2*k* *m*2),[[29]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-guo2006compression-29) where *k* is the number of edges to delete and *m* is the number of edges in the input graph.

**Matching**[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=10" \o "Edit section: Matching)]

A [matching](https://en.wikipedia.org/wiki/Matching_(graph_theory)) in a graph is a subset of its edges, no two of which share an endpoint. [Polynomial time](https://en.wikipedia.org/wiki/Polynomial_time) algorithms are known for many algorithmic problems on matchings, including [maximum matching](https://en.wikipedia.org/wiki/Maximum_matching) (finding a matching that uses as many edges as possible), [maximum weight matching](https://en.wikipedia.org/wiki/Maximum_weight_matching), and [stable marriage](https://en.wikipedia.org/wiki/Stable_marriage).[[30]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-30) In many cases, matching problems are simpler to solve on bipartite graphs than on non-bipartite graphs,[[31]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-31) and many matching algorithms such as the [Hopcroft–Karp algorithm](https://en.wikipedia.org/wiki/Hopcroft%E2%80%93Karp_algorithm) for maximum cardinality matching[[32]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-32) work correctly only on bipartite inputs.

As a simple example, suppose that a set {\displaystyle P} of people are all seeking jobs from among a set of {\displaystyle J} jobs, with not all people suitable for all jobs. This situation can be modeled as a bipartite graph {\displaystyle (P,J,E)}where an edge connects each job-seeker with each suitable job.[[33]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-33) A [perfect matching](https://en.wikipedia.org/wiki/Perfect_matching) describes a way of simultaneously satisfying all job-seekers and filling all jobs; [Hall's marriage theorem](https://en.wikipedia.org/wiki/Hall%27s_marriage_theorem) provides a characterization of the bipartite graphs which allow perfect matchings. The [National Resident Matching Program](https://en.wikipedia.org/wiki/National_Resident_Matching_Program) applies graph matching methods to solve this problem for [U.S. medical student](https://en.wikipedia.org/wiki/Medical_education_in_the_United_States) job-seekers and [hospital residency](https://en.wikipedia.org/wiki/Residency_(medicine)) jobs.[[34]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-34)

The [Dulmage–Mendelsohn decomposition](https://en.wikipedia.org/wiki/Dulmage%E2%80%93Mendelsohn_decomposition) is a structural decomposition of bipartite graphs that is useful in finding maximum matchings.[[35]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-35)

Additional applications[[edit](https://en.wikipedia.org/w/index.php?title=Bipartite_graph&action=edit&section=11" \o "Edit section: Additional applications)]

Bipartite graphs are extensively used in modern [coding theory](https://en.wikipedia.org/wiki/Coding_theory), especially to decode [codewords](https://en.wikipedia.org/wiki/Codeword) received from the channel. [Factor graphs](https://en.wikipedia.org/wiki/Factor_graph) and [Tanner graphs](https://en.wikipedia.org/wiki/Tanner_graph) are examples of this. A Tanner graph is a bipartite graph in which the vertices on one side of the bipartition represent digits of a codeword, and the vertices on the other side represent combinations of digits that are expected to sum to zero in a codeword without errors.[[36]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-36) A factor graph is a closely related [belief network](https://en.wikipedia.org/wiki/Belief_network) used for probabilistic decoding of [LDPC](https://en.wikipedia.org/wiki/LDPC) and [turbo codes](https://en.wikipedia.org/wiki/Turbo_codes).[[37]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-37)

In computer science, a [Petri net](https://en.wikipedia.org/wiki/Petri_net) is a mathematical modeling tool used in analysis and simulations of concurrent systems. A system is modeled as a bipartite directed graph with two sets of nodes: A set of "place" nodes that contain resources, and a set of "event" nodes which generate and/or consume resources. There are additional constraints on the nodes and edges that constrain the behavior of the system. Petri nets utilize the properties of bipartite directed graphs and other properties to allow mathematical proofs of the behavior of systems while also allowing easy implementation of simulations of the system.[[38]](https://en.wikipedia.org/wiki/Bipartite_graph#cite_note-38)

In [projective geometry](https://en.wikipedia.org/wiki/Projective_geometry), [Levi graphs](https://en.wikipedia.org/wiki/Levi_graph) are a form of bipartite graph used to model the incidences between points and lines in a [configuration](https://en.wikipedia.org/wiki/Configuration_(geometry)). Corresponding to the geometric property of points and lines that every two lines meet in at most one point and every two points be connected with a single line, Levi graphs necessarily do not contain any cycles of length four, so their [girth](https://en.wikipedia.org/wiki/Girth_(graph_theory)) must be six or more

### 拓扑排序

偏序=》全序

类似选择排序，每次选度为0的。

引申：计算入度、出度

### 单源路径

**final int**[][] **times**;  
 **final int K**;  
 **int N**;  
  
 HashSet<Integer> **arrived** = **new** HashSet<>();  
 **public int**[] **distance**;  
 *//经过的索引路径* **public** List<List<Integer>> **path** = **new** ArrayList<>();  
  
 **public** Dijkstra(**int**[][] times, **int** k) {  
 **this**.**times** = times;  
 **K** = k;  
 **N** = times[0].**length**;  
 **distance** = **new int**[**N**];  
 **for** (**int** i = 0; i < **N**; i++) {  
 **distance**[i] = times[**K**][i];  
 **path**.add(**new** ArrayList<>());  
 }  
 **arrived**.add(**K**);  
 }  
  
 **public void** run() {  
 *//每次加入一个点* **for** (**int** i = 0; i < **N**-1; i++) {  
 **int** nearIndex = getNear();  
 **if** (nearIndex == -1){  
 **return**;  
 }  
 **arrived**.add(nearIndex);  
 updateDistance(nearIndex);  
 }  
 }  
  
 **private void** updateDistance(**int** bridgeIndex){  
 **for** (**int** i = 0; i < **N**; i++) {  
 **if**(**arrived**.contains(i)){  
 **continue**;  
 }  
 **if** (**distance**[i] - **distance**[bridgeIndex]> **times**[bridgeIndex][i]){  
 *//加号越界  
// if (distance[i] > distance[bridgeIndex] + times[bridgeIndex][i]){* **distance**[i] = **distance**[bridgeIndex] + **times**[bridgeIndex][i];  
 List<Integer> thisPath = **path**.get(i);  
 thisPath.add(i);  
 }  
 }  
 }  
  
 **private int** getNear(){  
 **int** minDistance = Integer.***MAX\_VALUE***;  
 **int** minIndex = -1;  
 **for** (**int** j = 0; j < **N**; j++) {  
 **if** (**arrived**.contains(j)){  
 **continue**;  
 }  
 **if** (minDistance > **distance**[j]){  
 minDistance = **distance**[j];  
 minIndex = j;  
 }  
 }  
 **return** minIndex;  
 }

<https://leetcode.com/problems/cheapest-flights-within-k-stops/description/>

单源路径

743. Network Delay Time

带权单源路径

### 多源路径

带权多源路径

399. Evaluate Division：除式链接/路径查找

## 01背包问题

只能穷举

NP问题

字符串匹配

318. Maximum Product of Word Lengths

## 并查集

**int**[] **parent**;  
**int**[] **rank**;  
*//每个集合元素个数***int**[] **count**;  
*//集合总数***int countSet**;  
  
**public** DisjointCountSet(**int** n) {  
 **parent** = **new int**[n];  
 **rank** = **new int**[n];  
 **count** = **new int**[n];  
}  
  
*/\*\*  
 \* 要求本来没有。  
 \** ***@param x*** *\*/***public void** makeSet(**int** x) {  
 **parent**[x] = x;  
 **rank**[x] = 0;  
 **count**[x] = 1;  
 **countSet**++;  
}  
  
**public void** union(**int** x, **int** y) {  
 **if** (findSet(x) == findSet(y)){  
 **return**;  
 }  
 **countSet**--;  
 **int** cx = getCount(x),  
 cy = getCount(y);  
 setCount(x, cx+cy);  
 setCount(y, cx +cy);  
 link(findSet(x), findSet(y));  
}  
  
**private void** link(**int** set, **int** set1) {  
 **if** (**rank**[set] > **rank**[set1]) {  
 **parent**[set1] = set;  
  
 } **else if** (**rank**[set] < **rank**[set1]) {  
 **parent**[set] = set1;  
 } **else** {  
 **rank**[set1] += 1;  
 **parent**[set] = set1;  
 }  
  
}  
  
**public int** findSet(**int** x) {  
 **if** (**parent**[x] != x) {  
 **int** parentIndex = findSet(**parent**[x]);  
 **parent**[x] = parentIndex;  
 }  
 **return parent**[x];  
}  
  
**public int** getCountSet(){  
 **return countSet**;  
}  
  
**public int** getCount(**int** i){  
 **int** parent = findSet(i);  
 **return count**[parent];  
}  
  
**public void** setCount(**int** i, **int** c){  
 **int** parent = findSet(i);  
 **count**[parent] = c;  
}  
  
**public int** getMaxCount(){  
 **int** max = 0;  
 **for** (**int** i = 0; i < **count**.**length**; i++) {  
 **if** (max < **count**[i]){  
 max = **count**[i];  
 }  
 }  
 **return** max;  
}

200. Number of Islands

## 字符串匹配

朴素

Rabin-Karp/RollingHash

自动机

KMP

未整理

几个定理

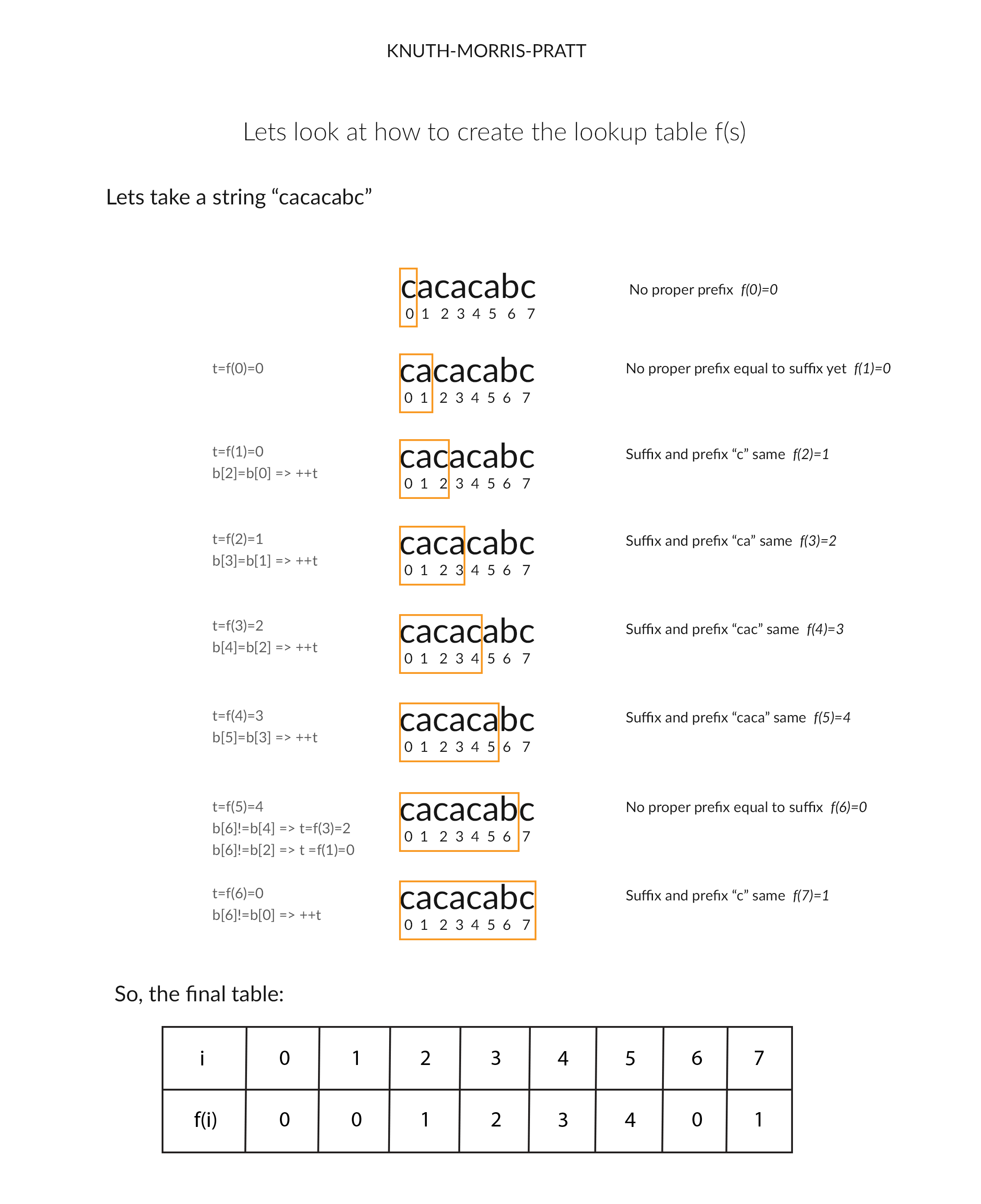
对每个模式串 P 而言，都有一个相应的一个匹配自动机。如图给出了一个模式 P = ababaca 的自动机构造过程：



为了构造一个自动机，我们应当首先定义一个辅助函数σ，称为相应模式串 P[1..m] 的后缀函数。函数 σ 是一个从 ∑\* 到 {0, 1, 2, ..., m} 上定义的映射。σ(x)表示文本串 x 的后缀的长度，且该后缀是模式串 P 的最长前缀。  
即 σ(x) = max{k : Pk ⊐ x}。例如，对于模式串 P = ab，有 σ(ɛ) = 0, σ(ccaca) = 1, σ(ccab) = 2。对于一个长度为 m 的模式串 P 而言，当且仅当 P ⊐ x 时，σ(x) = m。根据后缀函数的定义有：x ⊐ y，则 σ(x) ≤ σ(y)。

所以，对于给定模式串 P[1..m]，对应字符串匹配自动机定义如下：  
　　状态集 Q = {0, 1, ..., m}，初始状态 q0 = 0，接受状态 A = {m}；  
　　对任意状态 q 和字符 a，变迁函数 δ 定义为：δ(q, a) = σ(Pqa)。

我们之所以有 δ(q, a) = σ(Pqa)，是为了追踪当前已匹配最长的模式串 P 的前缀。考虑当前读取的最后一个字符 T[i]，为了寻找文本串 T 的一个子字符串可以匹配模式串 P 的前缀 Pj，Pj一定是 T[i] 的后缀。  
假设状态 q 是读取 T[i] 后的状态，即 q = φ(T[i])。因为我们有变迁函数δ，所以当其状态 q 能够告诉我们匹配 T[i] 后缀的P的最长前缀的长度，即在状态 q 下，Pq ⊐ Ti 且 q = σ(Ti)。  
所以当 q = m 时，便可以知道匹配成功。因为 φ(Ti) 和 σ(Ti) 都等于 q，所以自动机运行时能够保持如下不变式：φ(Ti) = σ(Ti)。



## 计算几何

### 寻找凸包

1. Graham扫描法
2. Jarvis步进法

### 寻找最近点对

## 扩展

### 马拉车算法

5. Longest Palindromic Substring

**1.Manacher算法原理与实现**

下面介绍Manacher算法的原理与步骤。

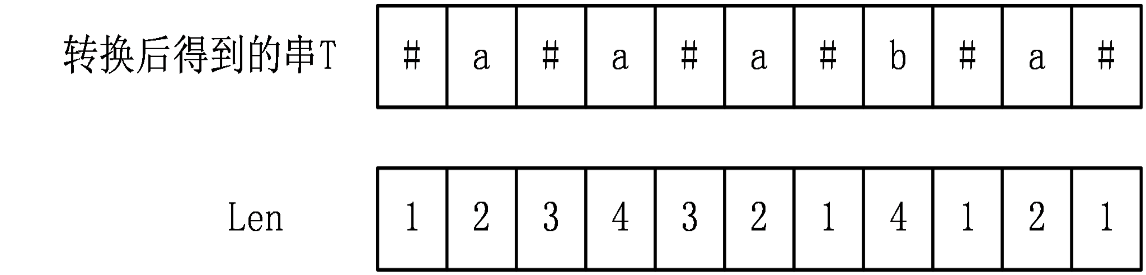
首先，Manacher算法提供了一种巧妙地办法，将长度为奇数的回文串和长度为偶数的回文串一起考虑，具体做法是，在原字符串的每个相邻两个字符中间插入一个分隔符，同时在首尾也要添加一个分隔符，分隔符的要求是不在原串中出现，一般情况下可以用#号。下面举一个例子：



**（1）Len数组简介与性质**

Manacher算法用一个辅助数组Len[i]表示以字符T[i]为中心的最长回文字串的最右字符到T[i]的长度，比如以T[i]为中心的最长回文字串是T[l,r],那么Len[i]=r-i+1。

对于上面的例子，可以得出Len[i]数组为:



Len数组有一个性质，那就是Len[i]-1就是该回文子串在原字符串S中的长度，至于证明，首先在转换得到的字符串T中，所有的回文字串的长度都为奇数，那么对于以T[i]为中心的最长回文字串，其长度就为2\*Len[i]-1,经过观察可知，T中所有的回文子串，其中分隔符的数量一定比其他字符的数量多1，也就是有Len[i]个分隔符，剩下Len[i]-1个字符来自原字符串，所以该回文串在原字符串中的长度就为Len[i]-1。

有了这个性质，那么原问题就转化为求所有的Len[i]。下面介绍如何在线性时间复杂度内求出所有的Len。

**（2）Len数组的计算**

首先从左往右依次计算Len[i]，当计算Len[i]时，Len[j](0<=j<i)已经计算完毕。设P为之前计算中最长回文子串的右端点的最大值，并且设取得这个最大值的位置为po，分两种情况：

第一种情况：i<=P

那么找到i相对于po的对称位置，设为j，那么如果Len[j]<P-i，如下图：



那么说明以j为中心的回文串一定在以po为中心的回文串的内部，且j和i关于位置po对称，由回文串的定义可知，一个回文串反过来还是一个回文串，所以以i为中心的回文串的长度至少和以j为中心的回文串一样，即Len[i]>=Len[j]。因为Len[j]<P-i,所以说i+Len[j]<P。由对称性可知Len[i]=Len[j]。

如果Len[j]>=P-i,由对称性，说明以i为中心的回文串可能会延伸到P之外，而大于P的部分我们还没有进行匹配，所以要从P+1位置开始一个一个进行匹配，直到发生失配，从而更新P和对应的po以及Len[i]。



第二种情况: i>P

如果i比P还要大，说明对于中点为i的回文串还一点都没有匹配，这个时候，就只能老老实实地一个一个匹配了，匹配完成后要更新P的位置和对应的po以及Len[i]。



**2.时间复杂度分析**

Manacher算法的时间复杂度分析和Z算法类似，因为算法只有遇到还没有匹配的位置时才进行匹配，已经匹配过的位置不再进行匹配，所以对于T字符串中的每一个位置，只进行一次匹配，所以Manacher算法的总体时间复杂度为O(n)，其中n为T字符串的长度，由于T的长度事实上是S的两倍，所以时间复杂度依然是线性的。

5. Longest Palindromic Substring

**final** String **src**;  
  
**char**[] **extend**;  
**int**[] **halfWidth**;  
  
**public** List<String> **longest** = **new** ArrayList<>();  
  
**public** Manacher(String src) {  
 **this**.**src** = src;  
 **extend** = **new char**[2\* src.length()+1];  
 **halfWidth** = **new int**[2\*src.length()+1];  
}  
  
**public void** run(){  
 insertBoundary();  
 computeEveryLength();  
 computeLongest();  
}  
  
**void** computeLongest(){  
 **int** max = 0;  
 List<Integer> maxIndexes = **new** ArrayList<>();  
 **for** (**int** i = 0; i < **src**.length()\*2+1; i++) {  
 **if** (**halfWidth**[i] < max){  
 **continue**;  
 }  
 **if** (**halfWidth**[i] == max){  
 maxIndexes.add(i);  
 **continue**;  
 }  
 maxIndexes.clear();  
 maxIndexes.add(i);  
 max = **halfWidth**[i];  
 }  
 maxIndexes.stream().forEach(index->{  
 StringBuilder stringBuilder = **new** StringBuilder();  
 **for** (**int** i = index - **halfWidth**[index] + 1; i < index + **halfWidth**[index]; i++) {  
 **if** (**extend**[i] != **'#'**){  
 stringBuilder.append(**extend**[i]);  
 }  
 }  
 **longest**.add(stringBuilder.toString());  
 });  
}  
  
**void** computeEveryLength(){  
 **halfWidth**[0] = 1;  
 **int** rightEdge = 0, center = 0;  
 **for** (**int** i = 1; i < **src**.length()\*2+1; i++) {  
 **int** thisHalfWidth = 0;  
 **if** (i <= rightEdge){  
 *//对称位置* thisHalfWidth = **halfWidth**[center-(i-center)];  
 **if** (i + thisHalfWidth-1 < rightEdge){  
 *//在已探测范围内，必然无法拓展* **halfWidth**[i] = thisHalfWidth;  
 **continue**;  
 }  
 thisHalfWidth = rightEdge - i + 1;  
 }**else**{  
 thisHalfWidth = 1;  
 }  
 *//以i为中心扩展* **while**(i + thisHalfWidth< **src**.length()\*2+1  
 && i - thisHalfWidth >= 0  
 && **extend**[i + thisHalfWidth] == **extend**[i - thisHalfWidth]){  
 thisHalfWidth++;  
 }  
 **halfWidth**[i] = thisHalfWidth;  
 **if** (i + thisHalfWidth-1 > rightEdge){  
 *//探测范围已拓展* rightEdge = i+ thisHalfWidth-1;  
 center = i;  
 }  
 }  
}  
  
**void** insertBoundary(){  
 **for** (**int** i = 0; i < **src**.length(); i++) {  
 **extend**[2\*i] = **'#'**;  
 **extend**[2\*i+1] = **src**.charAt(i);  
 }  
 **extend**[**src**.length() \*2] = **'#'**;  
}

### 格雷码/循环码/最小差错码

N位格雷码

for (int i = 0; i < 1<<n; i++)

result.add(i ^ i>>1);

如何生成：

1 递归

2 2进制公式//TODO

3 卡诺图

3 异或乘除

### 四平方定理/三平方定理

常规方法：

for(i:1-n) for(j:1-i) diff = i j\*j;

定理：

a natural number can be represented as the sum of three squares of integers

{\displaystyle n=x^{2}+y^{2}+z^{2}} n=x^{2}+y^{2}+z^{2}

if and only if n is not of the form {\displaystyle n=4^{a}(8b+7)} n = 4^a(8b + 7) for integers a and b.

注意x、y、z可以是0.

int numSquares(int n) {

//去除公式中的4

while ((n & 3) == 0) //n%4 == 0

n >>= 2;

if ((n & 7) == 7) return 4; //n % 8 == 7

if(is\_square(n)) return 1;

int sqrt\_n = (int) sqrt(n);

for(int i = 1; i<= sqrt\_n; i++){

if (is\_square(n-i\*i)) return 2;

}

return 3;

}

**int** **is\_square**(**int** n){

**int** temp = (**int**) sqrt(n);

**return** temp \* temp == n;

}

## 简单

### 逆波兰表达式

150. Evaluate Reverse Polish Notation

# 从形式到思路

## 广义表

341. Flatten Nested List Iterator

385. Mini Parser

394. Decode String

借助栈实现

## 数字

### 构建数字

321. Create Maximum Number

Given two arrays of length m and n with digits 0-9 representing two numbers. Create the maximum number of length k <= m + nfrom digits of the two. The relative order of the digits from the same array must be preserved. Return an array of the k digits.

高位尽量大=》贪心选取=》分析选取范围

402. Remove K Digits：remove *k* digits from the number so that the new number is the smallest possible.

解法一：

删除数字-》保留数字-》高位重要

解法二：

one can simply scan from left to right, and remove the first "peak" digit; the peak digit is larger than its right neighbor. One can repeat this procedure k times. One can simulate the above procedure by using a stack, and obtain a O(n) algorithm. Note, when the result stack (i.e. res) pop a digit, it is equivalent as remove that "peak" digit.

使用栈

738. Monotone Increasing Digits

when the number is 123454321, we could have a candidate of 123449999. It seems like a decent strategy is to take a monotone increasing prefix of N, then decrease the number before the "cliff" (the index where adjacent digits decrease for the first time) if it exists, and replace the rest of the characters with 9s.

例外123444321

### 位操作

397. Integer Replacement

If n is even, replace n with n/2.

If n is odd, you can replace n with either n + 1 or n - 1.

What is the minimum number of replacements needed for n to become 1?

S:

If n is even, halve it.

If n=3 or n-1 has less 1's than n+1, decrement n.

Otherwise, increment n.

doing bitCount on every iteration is not the best way. It is enough to examine the last two digits to figure out whether incrementing or decrementing will give more 1's. Indeed, if a number ends with 01, then certainly decrementing is the way to go. Otherwise, if it ends with 11, then certainly incrementing is at least as good as decrementing (\*011 -> \*010 / \*100) or even better (if there are three or more 1's).

海明距离

477. Total Hamming Distance：数字集合的总海明距离

按未划分成两组，当前位的贡献k\*(n-k).

421. Maximum XOR of Two Numbers in an Array

整体考虑：

1 从高位到低位

2 a^b = c <=> c^b=a

3 对于每一位，对于结果xx1yy, 逆转一下，查找元素。xx1 ^ b = a（b是已知元素，a是待查找元素）, 判断a存在

4 hash

338. Counting Bits：n个数

dp 后面的数依赖前面+1

779. K-th Symbol in Grammar

画图找关系

**return** Integer.bitCount(K-1) & 1;

***Updates***: (all index discussed below are 0-based)  
***Observation 1***: **N** does not matter as long as "K will be an integer in the range [1, 2^(N-1)]". We can ignore **N**.

***Observation 2***: let f(k) be the value of kth position (0-based), then:  
f(2 \* k) = 0 {if f(k) = 0} or, 1 {if f(k) = 1} => f(2 \* k) = f(k) xor 0  
f(2 \* k + 1) = 0 {if f(k) = 1} or 1 {if f(k) = 0} => f(2 \* k + 1) = f(k) xor 1

***Obervation 3***: if binary string of **k** is used, let **k = 1001010**, then we have:  
f(1001010) = f(100101) ^ 0 = f(10010) ^ 1 ^ 0 = f(1001) ^ 0 ^ 1 ^ 0 = ... = f(0) ^ 1 ^ 0 ^ 0 ^1 ^ 0 ^ 1 ^ 0 = 1 ^ 0 ^ 0 ^1 ^ 0 ^ 1 ^ 0  
So, the result is the **xor** operation on all bits of **k**. Since 0 does not change **xor** result, we can ignore all 0s.  
f(1001010) = 1 ^ 1 ^ 1 = (1^1) ^ 1 = 0 ^ 1 = 1  
f(11110011) = 1 ^ 1^ 1 ^ 1 ^ 1 ^1 = (1 ^ 1) ^ (1 ^ 1) ^ (1 ^1) = 0  
Now, it's easy to tell f(k) = 0 if **k** has **even** number of 1s in binary representation, and f(k) = 1 when **k** has **odd** number of 1s

## 从回溯/递归/推进到dp

494. Target Sum：n个数，1个和，一共有多少种方式。

//TODO 推出dp方案

## 数组/矩阵区间和/积

1. 303. Range Sum Query - Immutable

累积和

304. Range Sum Query 2D – Immutable

dp/累积和 //TODO 推算一下DP细节



307. Range Sum Query – Mutable

308．Range Sum Query 2D – Mutable

2维树状数组/线段树

## 四则运算

241. Different Ways to Add Parentheses

Given a string of numbers and operators, return all possible results

((2-1)-1) = 0

(2-(1-1)) = 2

穷举，分治，备忘录

399. Evaluate Division：连环计算

Given a / b = 2.0, b / c = 3.0.   
queries are: a / c = ?

构造图搜索

## 字符/数字集合

通常用数组来存储，但是没有前后关系，跟索引也没有关系。

### 相同字母异序词Anagram

字母数相等，与位置无关。

### 排列/组合

357. Count Numbers with Unique Digits：

Given a **non-negative** integer n, count all numbers with unique digits, x, where 0 ≤ x < 10n.

**Example:**  
Given n = 2, return 91. (The answer should be the total numbers in the range of 0 ≤ x < 100, excluding [11,22,33,44,55,66,77,88,99])

60. Permutation Sequence：Given *n* and *k*, return the *k*th permutation sequence.

按规律推导

46. Permutations：Given a collection of **distinct** integers, return all possible permutations.

47. Permutations II

重复的下一个排列

556. Next Greater Element III：Given a positive **32-bit** integer **n**, you need to find the smallest **32-bit** integer which has exactly the same digits existing in the integer **n**and is greater in value than n. If no such positive **32-bit** integer exists, you need to return -1.

77. Combinations

递归构造。

752. Open the Lock

两端bfs

### 元素和

377. Combination Sum IV：Given an integer array with all positive numbers and no duplicates, find the number of possible combinations that add up to a positive integer target.

***nums*** = [1, 2, 3]

***target*** = 4

The possible combination ways are:

(1, 1, 1, 1)、(1, 1, 2)。。。

回溯-》dp

### 背包问题/选或不选

集合划分问题

https://en.wikipedia.org/wiki/Partition\_problem

416. Partition Equal Subset Sum

698. Partition to K Equal Sum Subsets

不同角度，穷尽搜索

<https://leetcode.com/problems/partition-to-k-equal-sum-subsets/solution/>

Approach #1: Search by Constructing Subset Sums

Approach #2: Dynamic Programming on Subsets of Input

473. Matchsticks to Square

Dfs 大值在前

背包问题

https://leetcode.com/problems/ones-and-zeroes/description/

39. Combination Sum

选用多次

40. Combination Sum II

集合中有重复元素，选用一次

216. Combination Sum III

元素不可重复

377. Combination Sum IV

### 推进

<https://leetcode.com/problems/house-robber-ii/description/>

DP

## 字符串

字符间有位置关系

### 字符串匹配

187. Repeated DNA Sequences

All DNA is composed of a series of nucleotides abbreviated as A, C, G, and T, for example: "ACGAATTCCG". When studying DNA, it is sometimes useful to identify repeated sequences within the DNA.

Write a function to find all the 10-letter-long sequences (substrings) that occur more than once in a DNA molecule.

Rollinghash把字符串映射成数字，再比较

### 两个字符串的关系

583. Delete Operation for Two Strings: Given two words *word1* and *word2*, find the minimum number of steps required to make *word1* and *word2* the same, where in each step you can delete one character in either string.

解1：最长公共子串：比较第一个字母；递归+备忘=》dp

缩小规模-》DP

https://leetcode.com/problems/minimum-swaps-to-make-sequences-increasing/description/

722. Remove Comments

line comments, and block comments.

注意嵌套情况

“”字符串嵌套，转义字符

//可以在一行的任意位置

/\*/不完整

https://leetcode.com/problems/unique-substrings-in-wraparound-string/description/

https://leetcode.com/problems/unique-substrings-in-wraparound-string/discuss/95439/Concise-Java-solution-using-DP

712. Minimum ASCII Delete Sum for Two Strings

### 拾遗

3. Longest Substring Without Repeating Characters：Given a string, find the length of the **longest substring** without repeating characters.

解：If a substring s\_{ij}*s*​*ij*​​ from index i*i* to j - 1*j*−1 is already checked to have no duplicate characters. We only need to check if s[j]*s*[*j*] is already in the substring s\_{ij}*s*​*ij*​​.

Back to our problem. We use HashSet to store the characters in current window [i, j)[*i*,*j*) (j = i*j*=*i* initially). Then we slide the index j*j* to the right. If it is not in the HashSet, we slide j*j* further. Doing so until s[j] is already in the HashSet. At this point, we found the maximum size of substrings without duplicate characters start with index i*i*. If we do this for all i*i*, we get our answer.

## 回文

回文利用位置对称关系

1. 中心拓展
2. 马拉车

子串

5. Longest Palindromic Substring

马拉车算法

647. Palindromic Substrings：一个字符串有多少个回文子串

如上章节总结所述

516. Longest Palindromic Subsequence

dp, 是否要左右端点的字母

131. Palindrome Partitioning

Given a string *s*, partition *s* such that every substring of the partition is a palindrome.

Return all possible palindrome partitioning of *s*.

找到第一个回文，然后递归余下的

132. Palindrome Partitioning II：Return the minimum cuts needed for a palindrome partitioning of *s*.

递归+备忘=dp

子序列（中漏）

https://leetcode.com/problems/longest-palindromic-subsequence/description/

dp[i][j]: the longest palindromic subsequence's length of substring(i, j)

State transition:

dp[i][j] = dp[i+1][j-1] + 2 if s.charAt(i) == s.charAt(j)

otherwise, dp[i][j] = Math.max(dp[i+1][j], dp[i][j-1])

Initialization: dp[i][i] = 1

214. Shortest Palindrome：Given a string **s**, find the longest palindromic substring in **s**.

解：转化为前缀后缀关系：we reserved the original string s*s* and stored it as \text{rev}rev. We iterate over i*i* from 00 to n-1*n*−1 and check for s[0:n-i] == rev[i:]*s*[0:*n*−*i*]==*rev*[*i*:]. Pondering over this statement, had the \text{rev}rev been concatenated to s*s*, this statement is just finding the longest prefix that is equal to the suffix.

正向的前缀 = 反向的后缀，剩余的就是需要添加的字母。

* We use the KMP lookup table generation
* Create new\_snew\_s as s+"#"+reverse(s)s+"#"+reverse(s) and use the string in the lookup-generation algorithm
  + The "#" in the middle is required, since without the #, the 2 strings could mix with each ther, producing wrong answer. For example, take the string \text{"aaaa"}"aaaa". Had we not inserted "#" in the middle, the new string would be \text{"aaaaaaaa"}"aaaaaaaa" and the largest prefix size would be 7 corresponding to "aaaaaaa" which would be obviously wrong. Hence, a delimiter is required at the middle.
* Return reversed string after the largest palindrome from beginning length(given by n−f[n\_new-1]n−f[n\_new-1]) + original string s*s*

336. Palindrome Pairs

Given a list of **unique** words, find all pairs of ***distinct*** indices (i, j) in the given list, so that the concatenation of the two words, i.e. words[i] + words[j] is a palindrome.

**Example 2:**  
Given words = ["abcd", "dcba", "lls", "s", "sssll"]  
Return [[0, 1], [1, 0], [3, 2], [2, 4]]  
The palindromes are ["dcbaabcd", "abcddcba", "slls", "llssssll"]

O(nk^2):字典树/map

## 单词列表

809. Expressive Words

Sometimes people repeat letters to represent extra feeling, such as "hello" -> "heeellooo", "hi" -> "hiiii".  Here, we have groups, of adjacent letters that are all the same character, and adjacent characters to the group are different.  A group is extended if that group is length 3 or more, so "e" and "o" would be extended in the first example, and "i" would be extended in the second example.  As another example, the groups of "abbcccaaaa" would be "a", "bb", "ccc", and "aaaa"; and "ccc" and "aaaa" are the extended groups of that string.

For some given string S, a query word is *stretchy* if it can be made to be equal to S by extending some groups.  Formally, we are allowed to repeatedly choose a group (as defined above) of characters c, and add some number of the same character c to it so that the length of the group is 3 or more.  Note that we cannot extend a group of size one like "h" to a group of size two like "hh" - all extensions must leave the group extended - ie., at least 3 characters long.

Given a list of query words, return the number of words that are stretchy.

**Example:**

**Input:**

S = "heeellooo"

words = ["hello", "hi", "helo"]

**Output:** 1

**Explanation:**

We can extend "e" and "o" in the word "hello" to get "heeellooo".

We can't extend "helo" to get "heeellooo" because the group "ll" is not extended.

For some word, write the head character of every group, and the count of that group. For example, for "abbcccddddaaaaa", we'll write the "key" of "abcda", and the "count" [1,2,3,4,5].

Let's see if a word is stretchy. Evidently, it needs to have the same key as S.

Now, let's say we have individual counts c1 = S.count[i] and c2 = word.count[i].

* If c1 < c2, then we can't make the ith group of word equal to the ith word of S by adding characters.
* If c1 >= 3, then we can add letters to the ith group of word to match the ith group of S, as the latter is *extended*.
* Else, if c1 < 3, then we must have c2 == c1 for the ith groups of word and S to match.

524. Longest Word in Dictionary through Deleting：Given a string and a string dictionary, find the longest string in the dictionary that can be formed by deleting some characters of the given string.

解：简单：逐个比较是否是子序列

720. Longest Word in Dictionary

Given a list of strings words representing an English Dictionary, find the longest word in words that can be built one character at a time by other words in words. If there is more than one possible answer, return the longest word with the smallest lexicographical order.

If there is no answer, return the empty string.

**Example 1:**

**Input:**

words = ["w","wo","wor","worl", "world"]

**Output:** "world"

**Explanation:**

The word "world" can be built one character at a time by "w", "wo", "wor", and "worl".

解：字典树

查找：字典树

212. Word Search II

318. Maximum Product of Word Lengths：two not share common letters

一个整数标记一个单词的字母含量。位操作快一点点。

126. Word Ladder II

单词变种最短路径

双向bfs

注意，没有答案的，树的高度只要一半就可以了。

新生的节点，不要重复已有的节点。但是同一层的可以重复。

变种后集合的hash查找，比逐个单词比较是否变种，要快很多。

676. Implement Magic Dictionary

For the method search, you'll be given a word, and judge whether if you modify exactly one character into another character in this word, the modified word is in the dictionary you just built.

字典树

521. Longest Uncommon Subsequence I

解：较长的串；如果长度相等，则任意一个；若完全相等，则没有

522. Longest Uncommon Subsequence II

O(n^2)找最长的几个分析。

Sort the strings in the reverse order. If there is not duplicates in the array, then the longest string is the answer.

But if there are duplicates, and if the longest string is not the answer, then we need to check other strings. But the smaller strings can be subsequence of the bigger strings.  
For this reason, we need to check if the string is a subsequence of all the strings bigger than itself. If it's not, that is the answer.

## 数组

### 递推关系

368. Largest Divisible Subset：Given a set of **distinct** positive integers, find the largest subset such that every pair (Si, Sj) of elements in this subset satisfies: Si % Sj = 0 or Sj % Si = 0.

O(n^2):排序后，dp，每增加一个数，在前面的基础上，找到最长的。

### 排序

324. Wiggle Sort II：大小元素交叉摆

快排的应用：划分出大小分组。大小元素的交叉摆放。

791. Custom Sort String

S and T are strings composed of lowercase letters. In S, no letter occurs more than once.

S was sorted in some custom order previously. We want to permute the characters of T so that they match the order that S was sorted. More specifically, if x occurs before y in S, then x should occur before y in the returned string.

Return any permutation of T (as a string) that satisfies this property.

**Example :**

**Input:**

S = "cba"

T = "abcd"

**Output:** "cbad"

**Explanation:**

"a", "b", "c" appear in S, so the order of "a", "b", "c" should be "c", "b", and "a".

Since "d" does not appear in S, it can be at any position in T. "dcba", "cdba", "cbda" are also valid outputs.

解：计数排序

### 前K

373. Find K Pairs with Smallest Sums

You are given two integer arrays nums1 and nums2 sorted in ascending order and an integer k.

Define a pair (u,v) which consists of one element from the first array and one element from the second array.

Find the k pairs (u1,v1),(u2,v2) ...(uk,vk) with the smallest sums.

简单，记录数组第一个元素，匹配每一个。然后推进。

215. Kth Largest Element in an Array

快排

692. Top K Frequent Words

简单，堆排序

### 波形

496. Next Greater Element I

You are given two arrays **(without duplicates)** nums1 and nums2 where nums1’s elements are subset of nums2. Find all the next greater numbers for nums1's elements in the corresponding places of nums2.

503. Next Greater Element II:循环数组

O(n):从后往前推算大一点元素，下坡入栈，上坡出栈。

376. Wiggle Subsequence：A sequence of numbers is called a wiggle sequence if the differences between successive numbers strictly alternate between positive and negative. Given a sequence of integers, return the length of the longest subsequence that is a wiggle sequence.

解一：贪心：波形简化

解二：记录最后一个点作为波峰、波谷的最有值，再根据前面所有推导下一个点。https://leetcode.com/problems/wiggle-subsequence/solution/

股票系列

309. Best Time to Buy and Sell Stock with Cooldown

Input: [1,2,3,0,2]

Output: 3

Explanation: transactions = [buy, sell, cooldown, buy, sell]

状态机 + dp

714. Best Time to Buy and Sell Stock with Transaction Fee

Input: prices = [1, 3, 2, 8, 4, 9], fee = 2

Output: 8

Explanation: The maximum profit can be achieved by:

Buying at prices[0] = 1

Selling at prices[3] = 8

Buying at prices[4] = 4

Selling at prices[5] = 9

The total profit is ((8 - 1) - 2) + ((9 - 4) - 2) = 8.

计算时，考虑fee即可

456. 132 Pattern：checks whether there is a 132 pattern in the list：不一定连续

区间中的值

Input: [3, 1, 4, 2]

Output: True

Explanation: There is a 132 pattern in the sequence: [1, 4, 2].

解一：O(n^2):对于每一个中间值，向左找到最小的值，向有找最大的值

解二：O(n^2):波形

### 划分

813. Largest Sum of Averages：We partition a row of numbers A into at most K adjacent (non-empty) groups, then our score is the sum of the average of each group. What is the largest score we can achieve?

1. 回溯 + memo

2. 二维dp

dp(n,k) = max (dp(x, k-1) + x到n的平均数), x从1到n-1

### 子序列长度

一个字符串有n!个子序列

s是不是t的子序列

public boolean isSubsequence(String s, String t) {

if (s.length() == 0) return true;

int indexS = 0, indexT = 0;

while (indexT < t.length()) {

if (t.charAt(indexT) == s.charAt(indexS)) {

indexS++;

if (indexS == s.length()) return true;

}

indexT++;

}

return false;

}

491. Increasing Subsequences：find all the different possible increasing subsequences of the given array

dfs

673. Number of Longest Increasing Subsequence

### 丑数

质因子只含235的数

263. Ugly Number: check whether a given number is an ugly number. Ugly numbers are **positive numbers** whose prime factors only include 2, 3, 5.

解：不断除235

264. Ugly Number II

三个倍数队列包含了所有的丑数，然后归并排序。

313. Super Ugly Number

同上

### 排序数组查找

二分查找变种

### 逆序

775. Global and Local Inversions

O(n): find in range 0 to i-2, see if there is a element larger than A[i]

### 两个数的关系

2/3/4 Sum （smaller、closet）

暂定一个值，排序查找另一个值

注意平均值关系，运用两个指针单向滑动

### 小区间

思路：计数排序、二分查找

例：

825. Friends Of Appropriate Ages

### 区间和/积

累积和：累积和特点：对于单一元素，单调递增。

209. Minimum Size Subarray Sum:find the minimal length of a contiguous subarray of which the sum ≥ s

O(nlogn) 累计和，递增=》二分搜索，分割元素查找=》扣除固定值=》在和上扣除

O(n)窗口滑动

713. Subarray Product Less Than K

log(∏​*i*​​*x*​*i*​​)=∑​*i*​​log*x*​*i*​​，转化同上

明显：列举所有：O(n3)

累积和：O(n2)

累积和：注意脑中有一个递增序列：二分查找(nlogn)

两个指针：某一个index，有条件提前结束查找：（n）

238. Product of Array Except Self

空间利用，利用输出数组

Dp空间的优化

### 利用索引

数组实际上是一个map 可以在原数组操作

769. Max Chunks To Make Sorted：Given an array arr that is a permutation of [0, 1, ..., arr.length - 1],分组排序后，全数组有序。如何分最多的组？

O(n)：利用数值和索引的关系

768. Max Chunks To Make Sorted II：去除索引对应条件

O(nlogn):排序后，建立值和索引关系，转化成上一个问题。

O(n): 索引无关：Use two arrays to store the left max and right min: Iterate through the array, each time all elements to the left are smaller (or equal) to all elements to the right, there is a new chunck.

565. Array Nesting：A zero-indexed array A of length N contains all integers from 0 to N-1. Find and return the longest length of set S, where S[i] = {A[i], A[A[i]], A[A[A[i]]], ... } subjected to the rule below.

O(n)：构成链表，查找最长环.记录已访问。利用原数组，可以减少空间。

442. Find All Duplicates in an Array: Given an array of integers, 1 ≤ a[i] ≤ n (n = size of array), some elements appear twice and others appear once.

O(n)：用Map记录已访问数据-》使用原数组减少空间-》映射空间大的话，最终可以恢复数组

274. H-Index：Given an array of citations (each citation is a non-negative integer) of a researcher, write a function to compute the researcher's h-index.

定义：A scientist has in Given an array of citations (each citation is a non-negative integer) of a researcher, write a function to compute the researcher's h-index. dex h if h of his/her N papers have at least h citations each, and the other N − h papers have no more than h citations each

O(n)：总文章数可能比较小，计数排序，利用索引，从后往前计算累计和。

275. H-Index II：Given an array of citations in ascending order

O(nlogn)二分搜索

526. Beautiful Arrangement

Suppose you have **N** integers from 1 to N. We define a beautiful arrangement as an array that is constructed by these **N** numbers successfully if one of the following is true for the ith position (1 <= i <= N) in this array:

1. The number at the ith position is divisible by **i**.
2. **i** is divisible by the number at the ith position.

Now given N, how many beautiful arrangements can you construct?

O(n!):暴力列举所有验证

O(k): k refers to the number of valid permutations:在构建过程中验证

667. Beautiful Arrangement II

Given two integers n and k, you need to construct a list which contains n different positive integers ranging from 1 to n and obeys the following requirement:   
Suppose this list is [a1, a2, a3, ... , an], then the list [|a1 - a2|, |a2 - a3|, |a3 - a4|, ... , |an-1 - an|] has exactly k distinct integers.

O (n):贪心构建：

When \text{k = n-1}k = n-1, a valid construction is \text{[1, n, 2, n-1, 3, n-2, ....]}[1, n, 2, n-1, 3, n-2, ....]. One way to see this is, we need to have a difference of \text{n-1}n-1, which means we need \text{1}1 and \text{n}n adjacent; then, we need a difference of \text{n-2}n-2, etc.

Also, when \text{k = 1}k = 1, a valid construction is \text{[1, 2, 3, ..., n]}[1, 2, 3, ..., n]. So we have a construction when \text{n-k}n-k is tiny, and when it is large. This leads to the idea that we can stitch together these two constructions: we can put \text{[1, 2, ..., n-k-1]}[1, 2, ..., n-k-1]first so that \text{n}n is effectively \text{k+1}k+1, and then finish the construction with the first \text{"k = n-1"}"k = n-1" method.

For example, when \text{n = 6}n = 6 and \text{k = 3}k = 3, we will construct the array as \text{[1, 2, 3, 6, 4, 5]}[1, 2, 3, 6, 4, 5]. This consists of two parts: a construction of \text{[1, 2]}[1, 2] and a construction of \text{[1, 4, 2, 3]}[1, 4, 2, 3] where every element had \text{2}2 added to it (i.e. \text{[3, 6, 4, 5]}[3, 6, 4, 5]).

### 环/并查集

并查集问题也可以用dfs/bfs来解决。

565. Array Nesting

547. Friend Circles

统计不同集个数

### 拾遗

220. Contains Duplicate III

find out whether there are two distinct indices i and j in the array such that the absolute difference between nums[i] and nums[j] is at most t and the absolute difference between i and j is at most k.

窗口，TreeSet(红黑树)，桶

380. Insert Delete GetRandom O(1)

位置关系

777. Swap Adjacent in LR String

去掉X，比较LR

396. Rotate Function

Given an array of integers A and let *n* to be its length.

Assume Bk to be an array obtained by rotating the array A *k* positions clock-wise, we define a "rotation function" F on A as follow:

F(k) = 0 \* Bk[0] + 1 \* Bk[1] + ... + (n-1) \* Bk[n-1].

Calculate the maximum value of F(0), F(1), ..., F(n-1).

dp:数学推导：f(k) = f(k-1) + sum – nBk[0]

740. Delete and Earn

Given an array nums of integers, you can perform operations on the array.

In each operation, you pick any nums[i] and delete it to earn nums[i] points. After, you must delete **every** element equal to nums[i] - 1 or nums[i] + 1.

解：You start with 0 points. Return the maximum number of points you can earn by applying such operations.

Because all numbers are positive, if we "take" a number (use it to score points), we might as well take all copies of it, since we've already erased all its neighbors. We could keep a count of each number so we know how many points taking a number is worth total.

Now let's investigate what happens when we add a new number X (plus copies) that is larger than all previous numbers. Naively, our answer would be the previous answer, plus the value of X - which can be solved with dynamic programming. However, this fails if our previous answer had a number taken that was adjacent to X.

Luckily, we can remedy this. Let's say we knew using, the value of our previous answer, and avoid, the value of our previous answer that doesn't use the previously largest value prev. Then we could compute new values of using and avoid appropriately.

* Time Complexity (Java): We performed a radix sort instead, so our complexity is O(N+W)*O*(*N*+*W*) where W*W* is the range of allowable values for nums[i].
* Space Complexity (Java): O(W)*O*(*W*), the size of our count.

## 矩阵

解题思路：回溯、dfs、bfs、dp

路径搜索，一般只依赖邻居，通常可以用dp解决

最简单的空间占用是n2

有些可以优化到n，如2向路径。跟依赖多少方向有关系。

有些可以优化到1

417. Pacific Atlantic Water Flow：中间高两边低，水往两边流

解：

1. Two Queue and add all the Pacific border to one queue; Atlantic border to another queue.
2. Keep a visited matrix for each queue. In the end, add the cell visited by two queue to the result.  
   BFS: Water flood from ocean to the cell. Since water can only flow from high/equal cell to low cell, add the neighboor cell with height larger or equal to current cell to the queue and mark as visited.

### 矩阵2向路径

62 Unique Paths：左上到右下有多少种路径排列组合问题哦：DRRRDRRR（m\*n）!/(m!\*n!)

63. Unique Paths II：Now consider if some obstacles are added to the grids. How many unique paths would there be?

DP

### 矩阵4向路径

矩阵邻居搜索

79. Word Search：Given a 2D board and a word, find if the word exists in the grid.

回溯

212. Word Search II：Given a 2D board and a list of words from the dictionary, find all words in the board.

回溯+字典树

576. Out of Boundary Paths

There is an **m** by **n** grid with a ball. Given the start coordinate **(i,j)** of the ball, you can move the ball to **adjacent** cell or cross the grid boundary in four directions (up, down, left, right). However, you can **at most** move **N** times. Find out the number of paths to move the ball out of grid boundary.

出界

Dp：上一步推下一步，用矩阵记录上一步的值

542. 01 Matrix：Given a matrix consists of 0 and 1, find the distance of the nearest 0 for each cell.

简单dp

529. Minesweeper

解：搜索：Search rules:

1. If click on a mine ('M'), mark it as 'X', stop further search.
2. If click on an empty cell ('E'), depends on how many surrounding mine:  
   2.1 Has surrounding mine(s), mark it with number of surrounding mine(s), stop further search.  
   2.2 No surrounding mine, mark it as 'B', continue search its 8 neighbors.

### 矩阵搜索

74. Search a 2D Matrix

* Integers in each row are sorted from left to right.
* The first integer of each row is greater than the last integer of the previous row.

二分搜索，mid的计算比较复杂

240. Search a 2D Matrix II

* Integers in each row are sorted in ascending from left to right.
* Integers in each column are sorted in ascending from top to bottom.

判断右上角，每次排除一行/列

### 三角形

120. Triangle

Given a triangle, find the minimum path sum from top to bottom. Each step you may move to adjacent numbers on the row below.

For example, given the following triangle

[

[**2**],

[**3**,4],

[6,**5**,7],

[4,**1**,8,3]

]

The minimum path sum from top to bottom is 11 (i.e., **2** + **3** + **5** + **1** = 11).

坐标变换、dp

756. Pyramid Transition Matrix

We are stacking blocks to form a pyramid. Each block has a color which is a one letter string, like `'Z'`.

For every block of color `C` we place not in the bottom row, we are placing it on top of a left block of color `A` and right block of color `B`. We are allowed to place the block there only if `(A, B, C)` is an allowed triple.

We start with a bottom row of bottom, represented as a single string. We also start with a list of allowed triples allowed. Each allowed triple is represented as a string of length 3.

Return true if we can build the pyramid all the way to the top, otherwise false.

**Example 1:**

**Input:** bottom = "XYZ", allowed = ["XYD", "YZE", "DEA", "FFF"]

**Output:** true

**Explanation:**

We can stack the pyramid like this:

A

/ \

D E

/ \ / \

X Y Z

This works because ('X', 'Y', 'D'), ('Y', 'Z', 'E'), and ('D', 'E', 'A') are allowed triples.

一行一行处理，逐渐往上堆

字符可以看作数字

### 区域面积

拆分维度，分别dp，再综合计算

221. Maximal Square：Given a 2D binary matrix filled with 0's and 1's, find the largest square containing only 1's and return its area.

Dp

764. Largest Plus Sign：最大十字架

Dp

If we knew the longest possible arm length L\_u, L\_l, L\_d, L\_r*L*​*u*​​,*L*​*l*​​,*L*​*d*​​,*L*​*r*​​ in each direction from a center, we could know the order \min(L\_u, L\_l, L\_d, L\_r)min(*L*​*u*​​,*L*​*l*​​,*L*​*d*​​,*L*​*r*​​) of a plus sign at that center. We could find these lengths separately using dynamic programming.

### 拾遗

789. Escape The Ghosts

You are playing a simplified Pacman game. You start at the point (0, 0), and your destination is (target[0], target[1]). There are several ghosts on the map, the i-th ghost starts at (ghosts[i][0], ghosts[i][1]).

Each turn, you and all ghosts simultaneously \*may\* move in one of 4 cardinal directions: north, east, west, or south, going from the previous point to a new point 1 unit of distance away.

You escape if and only if you can reach the target before any ghost reaches you (for any given moves the ghosts may take.)  If you reach any square (including the target) at the same time as a ghost, it doesn't count as an escape.

Return True if and only if it is possible to escape.

**Example 1:**

**Input:**

ghosts = [[1, 0], [0, 3]]

target = [0, 1]

**Output:** true

**Explanation:**

You can directly reach the destination (0, 1) at time 1, while the ghosts located at (1, 0) or (0, 3) have no way to catch up with you.

**Example 2:**

**Input:**

ghosts = [[1, 0]]

target = [2, 0]

**Output:** false

**Explanation:**

You need to reach the destination (2, 0), but the ghost at (1, 0) lies between you and the destination.

解法：如果能比鬼先到终点，就能赢。

## 链表

19. Remove Nth Node From End两个指针O(n)

138. Copy List with Random Pointer利用原链表指针相对位置关系=》指针关系

链表环检测

环入口

142. Linked List Cycle II

解: using two pointers, one of them one step at a time. another pointer each take two steps. Suppose the first meet at step k,the length of the Cycle is r. so..2k-k=nr,k=nr

Now, the distance between the start node of list and the start node of cycle is s. the distance between the start of list and the first meeting node is k(the pointer which wake one step at a time waked k steps).the distance between the start node of cycle and the first meeting node is m, so...s=k-m,

s=nr-m=(n-1)r+(r-m),here we takes n = 1..so,

**using one pointer start from the start node of list, another pointer start from the first meeting node, all of them wake one step at a time, the first time they meeting each other is the start of the cycle.**

## 区间维护

56 Merge Intervals排序以后按顺序合并

228. Summary Ranges：Given a sorted integer array without duplicates, return the summary of its ranges.

646. Maximum Length of Pair Chain

You are given n pairs of numbers. In every pair, the first number is always smaller than the second number.

Now, we define a pair (c, d) can follow another pair (a, b) if and only if b < c. Chain of pairs can be formed in this fashion.

Given a set of pairs, find the length longest chain which can be formed. You needn't use up all the given pairs. You can select pairs in any order.

452. Minimum Number of Arrows to Burst Balloons

非重叠区域计数

435. Non-overlapping Intervals：Given a collection of intervals, find the minimum number of intervals you need to remove to make the rest of the intervals non-overlapping.

课程时间安排

https://en.wikipedia.org/wiki/Interval\_scheduling#Interval\_Scheduling\_Maximization

## 二叉树

236. Lowest Common Ancestor of a Binary Tree

根后序遍历

117. Populating Next Right Pointers in Each Node II

层次遍历

114. Flatten Binary Tree to Linked List

652. Find Duplicate Subtrees

把树转化为标志id的思想。

Suppose we have a unique identifier for subtrees: two subtrees are the same if and only if they have the same id.

Then, for a node with left child id of x and right child id of y, (node.val, x, y) uniquely determines the tree.

623. Add One Row to Tree

### 路径

路径dfs

113. Path Sum II：Given a binary tree and a sum, find all root-to-leaf paths where each path's sum equals the given sum.

437. Path Sum III：单条路径任意点

优化算法：每个路径构建前缀和

124. Binary Tree Maximum Path Sum：任意点到点

### Level

129. Sum Root to Leaf Numbers：Find the total sum of all root-to-leaf numbers.

换一个维度，从下往上，计算层次和。

337. House Robber III：二叉树必须隔层抢

Dp

遍历

103. Binary Tree Zigzag Level Order Traversal：from left to right, then right to left for the next level

## 树

## 图

207 Course Schedule依赖问题：拓扑排序（需要手写）时间问题：贪心算法

缔结斯科拉

https://leetcode.com/problems/cheapest-flights-within-k-stops/description/

332. Reconstruct Itinerary

欧拉回路

<https://www.cnblogs.com/acxblog/p/7390301.html>

<https://blog.csdn.net/u011466175/article/details/18861415>

环检测

802. Find Eventual Safe States

### 二分图

Our goal is trying to use two colors to color the graph and see if there are any adjacent nodes having the same color.

Initialize a color[] array for each node. Here are three states for colors[] array:

-1: Haven't been colored.

0: Blue.

1: Red.

For each node,

If it hasn't been colored, use a color to color it. Then use the other color to color all its adjacent nodes (DFS).

If it has been colored, check if the current color is the same as the color that is going to be used to color it. (Please forgive my english... Hope you can understand it.)

## 猜大小

375. Guess Number Higher or Lower II：when you guess a particular number x, and you guess wrong, you pay $x. You win the game when you guess the number I picked.

解：For each number x in range[i~j]  
we do: result\_when\_pick\_x = x + **max**{DP([i~x-1]), DP([x+1, j])}  
--> *// the max means whenever you choose a number, the feedback is always bad and therefore leads you to a worse branch.*  
then we get DP([i~j]) = **min**{xi, ... ,xj}  
--> *// this min makes sure that you are minimizing your cost.*

## 极大极小

486. Predict the Winner

Given an array of scores that are non-negative integers. Player 1 picks one of the numbers from either end of the array followed by the player 2 and then player 1 and so on. Each time a player picks a number, that number will not be available for the next player. This continues until all the scores have been chosen. The player with the maximum score wins.

Given an array of scores, predict whether player 1 is the winner. You can assume each player plays to maximize his score.

//TODO 回溯-》DP

649. Dota2 Senate

In the world of Dota2, there are two parties: the Radiant and the Dire.

The Dota2 senate consists of senators coming from two parties. Now the senate wants to make a decision about a change in the Dota2 game. The voting for this change is a round-based procedure. In each round, each senator can exercise one of the two rights:

1. Ban one senator's right:   
   A senator can make another senator lose **all his rights** in this and all the following rounds.
2. Announce the victory:   
   If this senator found the senators who still have rights to vote are all from **the same party**, he can announce the victory and make the decision about the change in the game.

Given a string representing each senator's party belonging. The character 'R' and 'D' represent the Radiant party and the Dire party respectively. Then if there are n senators, the size of the given string will be n.

The round-based procedure starts from the first senator to the last senator in the given order. This procedure will last until the end of voting. All the senators who have lost their rights will be skipped during the procedure.

Suppose every senator is smart enough and will play the best strategy for his own party, you need to predict which party will finally announce the victory and make the change in the Dota2 game. The output should be Radiant or Dire.

贪心，分析

https://leetcode.com/problems/can-i-win/discuss/95277/Java-solution-using-HashMap-with-detailed-explanation

## 拾遗

284. Peeking Iterator

封装一层，组合模式，提前获取next并缓存

621. Task Scheduler

Input: tasks = ["A","A","A","B","B","B"], n = 2

Output: 8

Explanation: A -> B -> idle -> A -> B -> idle -> A -> B.

767. Reorganize String: Given a string S, check if the letters can be rearranged so that two characters that are adjacent to each other are not the same.

O(nlogA):同上，先对每个字母计数，再排列

406. Queue Reconstruction by Height

一个人的位置，由比他高的人决定。

从高到低确定位置。

393. UTF-8 Validation

A character in UTF8 can be from 1 to 4 bytes long, subjected to the following rules:

For 1-byte character, the first bit is a 0, followed by its unicode code.

For n-bytes character, the first n-bits are all one's, the n+1 bit is 0, followed by n-1 bytes with most significant 2 bits being 10.

This is how the UTF-8 encoding would work:

Char. number range | UTF-8 octet sequence

(hexadecimal) | (binary)

--------------------+---------------------------------------------

0000 0000-0000 007F | 0xxxxxxx

0000 0080-0000 07FF | 110xxxxx 10xxxxxx

0000 0800-0000 FFFF | 1110xxxx 10xxxxxx 10xxxxxx

0001 0000-0010 FFFF | 11110xxx 10xxxxxx 10xxxxxx 10xxxxxx

捞针

260. Single Number III

全部两次+两个一次

按位划分成两个数组，转化成低级问题。

//more

# 常用解题思路

## 答案空间穷举回溯搜索

大多数问题都可以用该方法解决。有些问题存在效率更高的方式；有些不存在，例如背包问题。

需要有能力看出哪些问题不能优化。

参考资料，算法导论NP问题。

细节优化：可以采用双端bfs、减支方法优化。虽然不能提高O，可以快一点。

## DP

能够应用DP的问题，通常有两种可能

1. 可以构建无后效性递归式
2. 可以应用回溯算法，并且有大量重复。

表面上是TopDown算法，一般可以转化为bottomUp算法。

1. 队列推进，可以直接看出dp

DP的空间，

思考时，可以用较大的空间。

最终的优化，看递归式，到底依赖了多少上一步的结果。

https://leetcode.com/problems/2-keys-keyboard/discuss/105932/Java-solutions-from-naive-DP-to-optimized-DP-to-non-DP

常用解决方案：分治、二分、DP哈希、双指针、排序

双向bfs

贪心算法专题

关键是需要证明为什么贪心是对的。

狭义贪心算法是动态规划的特例。贪心算法代码看起来比动态规划简单，实际上思维上要复杂一点。算法导论有介绍321 Create Maximum Number广义贪心算法

回溯

心中一颗答案空间搜索树

两种搜索：dfs、bfs

Dfs用递归、栈实现

Bfs队列实现

# 常用正则表达式

# 测试细节

解题流程

确认算法、复杂度

确认输入边界，说明分支处理。

大于还是大于等于？

编写常规流程

代码走查

完善测试用例，确认测试通过。

Comparator注意越界

## 数字验证/解析

<https://leetcode.com/problems/valid-number/discuss/23977/A-clean-design-solution-By-using-design-pattern>

测试用例总结parsedouble的代码没有开源

整数、小数、科学计数法1. 前后空格2. 前0，多个03. 负04. 小数，后多个05. 小数，小数点前后没有06. 指数，指数为07. 字母、空格，中间夹杂8. 空值Pattern.matches("(\\+|-)?(\\d+(\\.\\d\*)?|\\.\\d+)(e(\\+|-)?\\d+)?", s);test(1, "123", true);test(2, " 123 ", true);test(3, "0", true);test(4, "0123", true); //Cannot agreetest(5, "00", true); //Cannot agreetest(6, "-10", true);test(7, "-0", true);test(8, "123.5", true);test(9, "123.000000", true);test(10, "-500.777", true);test(11, "0.0000001", true);test(12, "0.00000", true);test(13, "0.", true); //Cannot be more disagree!!!test(14, "00.5", true); /ly cannot agreetest(15, "123e1", true);test(16, "1.23e10", true);test(17, "0.5e-10", true);test(18, "1.0e4.5", false);test(19, "0.5e04", true);test(20, "12 3", false);test(21, "1a3", false);test(22, "", false);test(23, " ", false);test(24, null, false);test(25, ".1", true); //Ok, if you say sotest(26, ".", false);test(27, "2e0", true); //Really?!test(28, "+.8", true); (29, " 005047e+6", true); //Damn = =|||Here is the final Regex I got based on their definitionPattern.matches("(\\+|-)?(\\d+(\\.\\d\*)?|\\.\\d+)(e(\\+|-)?\\d+)?", s);But I thought my original one should be more rigorous!Pattern.matches("-?(([1-9]{1}+\\d\*|0)(\\.\\d+)?|\\.\\d+)(e-?[1-9]{1}+\\d\*)?", s);

数字解析

https://leetcode.com/problems/string-to-integer-atoi/discuss/4654/My-simple-solution

I think we only need to handle four cases:

discards all leading whitespaces

sign of the number

overflow

invalid input

过程中计算负数，因为负数范围大。

注意越界的处理

//容错比较低，需要提前验证

public static int parseInt(String s, int radix)

throws NumberFormatException

{

/\*

\* WARNING: This method may be invoked early during VM initialization

\* before IntegerCache is initialized. Care must be taken to not use

\* the valueOf method.

\*/

if (s == null) {

throw new NumberFormatException("null");

}

if (radix < Character.MIN\_RADIX) {

throw new NumberFormatException("radix " + radix +

" less than Character.MIN\_RADIX");

}

if (radix > Character.MAX\_RADIX) {

throw new NumberFormatException("radix " + radix +

" greater than Character.MAX\_RADIX");

}

int result = 0;

boolean negative = false;

int i = 0, len = s.length();

int limit = -Integer.MAX\_VALUE;

int multmin;

int digit;

if (len > 0) {

char firstChar = s.charAt(0);

if (firstChar < '0') { // Possible leading "+" or "-"

if (firstChar == '-') {

negative = true;

limit = Integer.MIN\_VALUE;

} else if (firstChar != '+')

throw NumberFormatException.forInputString(s);

if (len == 1) // Cannot have lone "+" or "-"

throw NumberFormatException.forInputString(s);

i++;

}

multmin = limit / radix;

while (i < len) {

// Accumulating negatively avoids surprises near MAX\_VALUE

digit = Character.digit(s.charAt(i++),radix);

if (digit < 0) {

throw NumberFormatException.forInputString(s);

}

//是否可以乘，升位

if (result < multmin) {

throw NumberFormatException.forInputString(s);

}

result \*= radix;

//是否可以增加本位的值

if (result < limit + digit) {

throw NumberFormatException.forInputString(s);

}

result -= digit;

}

} else {

throw NumberFormatException.forInputString(s);

}

return negative ? result : -result;

}

## 拾遗

测试

https://leetcode.com/problems/simplify-path/description/

* Did you consider the case where **path** = "/../"?  
  In this case, you should return "/".
* Another corner case is the path might contain multiple slashes '/' together, such as "/home//foo/".  
  In this case, you should ignore redundant slashes and return "/home/foo".

# 数学

## 数字

357. Count Numbers with Unique Digits

Given a **non-negative** integer n, count all numbers with unique digits, x, where 0 ≤ x < 10n.

**Example:**  
Given n = 2, return 91. (The answer should be the total numbers in the range of 0 ≤ x < 100, excluding [11,22,33,44,55,66,77,88,99])

解：超过10位的必定重复。

343. Integer Break：Given a positive integer n, break it into the sum of at least two positive integers and maximize the product of those integers.

The first thing we should consider is : What is the max product if we break a number N into two factors?

I use a function to express this product: f=x(N-x)

When x=N/2, we get the maximum of this function.

However, factors should be integers. Thus the maximum is (N/2)\*(N/2) when N is even or (N-1)/2 \*(N+1)/2 when N is odd.

When the maximum of f is larger than N, we should do the break.

(N/2)\*(N/2)>=N, then N>=4

(N-1)/2 \*(N+1)/2>=N, then N>=5

These two expressions mean that factors should be less than 4, otherwise we can do the break and get a better product. The factors in last result should be 1, 2 or 3. Obviously, 1 should be abandoned. Thus, the factors of the perfect product should be 2 or 3.

The reason why we should use 3 as many as possible is

For 6, 3 \* 3>2 \* 2 \* 2. Thus, the optimal product should contain no more than three 2.

Below is my accepted, O(N) solution.

**public** **class** Solution {

**public** int integerBreak(int n) {

**if**(n==2) **return** 1;

**if**(n==3) **return** 2;

int product = 1;

**while**(n>4){

product\*=3;

n-=3;

}

product\*=n;

**return** product;

}

}

## 运算符

### 乘

43. Multiply Strings: Given two non-negative integers num1 and num2 represented as strings, return the product of num1 and num2, also represented as a string.

### 除

29. Divide Two Integers：divide two integers without using multiplication, division and mod operator.

s1：位操作，变大除数s2：长除法

### 取模运算

定义

给定一个正整数p，任意一个整数n，一定存在等式 ：

n = kp + r ；

其中 k、r 是整数，且 0 ≤ r < p，则称 k 为 n 除以 p 的商，r 为 n 除以 p 的余数。

对于正整数 p 和整数 a,b，定义如下运算：

取模运算：a % p（或a mod p），表示a除以p的余数。

模p加法： ，其结果是a+b算术和除以p的余数。

模p减法： ，其结果是a-b算术差除以p的余数。

模p乘法： ，其结果是 a \* b算术乘法除以p的余数。

说明：

1. 同余式：正整数a，b对p取模，它们的余数相同，记做 或者a ≡ b (mod p)。

2. n % p 得到结果的正负由被除数n决定,与p无关。例如：7%4 = 3， -7%4 = -3， 7%-4 = 3， -7%-4 = -3。

基本性质

1. 若p|(a-b)，则a≡b (% p)。例如 11 ≡ 4 (% 7)， 18 ≡ 4(% 7)
2. (a % p)=(b % p)意味a≡b (% p)
3. 对称性：a≡b (% p)等价于b≡a (% p)
4. 传递性：若a≡b (% p)且b≡c (% p) ，则a≡c (% p)

运算规则

模运算与基本四则运算有些相似，但是除法例外。其规则如下：

1. (a + b) % p = (a % p + b % p) % p （1）
2. (a - b) % p = (a % p - b % p) % p （2）
3. (a \* b) % p = (a % p \* b % p) % p （3）
4. a ^ b % p = ((a % p)^b) % p （4）

* 结合律：

((a+b) % p + c) % p = (a + (b+c) % p) % p （5）

((a\*b) % p \* c)% p = (a \* (b\*c) % p) % p （6）

* 交换律：

(a + b) % p = (b+a) % p （7）

(a \* b) % p = (b \* a) % p （8）

* 分配律：

(a+b) % p = ( a % p + b % p ) % p （9）

((a +b)% p \* c) % p = ((a \* c) % p + (b \* c) % p) % p （10）

重要定理

* 若a≡b (% p)，则对于任意的c，都有(a + c) ≡ (b + c) (%p)；（11）
* 若a≡b (% p)，则对于任意的c，都有(a \* c) ≡ (b \* c) (%p)；（12）
* 若a≡b (% p)，c≡d (% p)，则 (a + c) ≡ (b + d) (%p)，(a - c) ≡ (b - d) (%p)，

(a \* c) ≡ (b \* d) (%p)，(a / c) ≡ (b / d) (%p)； （13）

应用

[编辑](javascript:;)

判别奇偶数

奇偶数的判别是模运算最基本的应用，也非常简单。

已知一个整数n对2取模，如果[余数](https://baike.baidu.com/item/%E4%BD%99%E6%95%B0)为0，则表示n为偶数，否则n为奇数。

C++实现功能函数：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10 | /\*  函数名：IsEven  函数功能：判别整数n的奇偶性。能被2整除为偶数，否则为奇数  输入值：intn，整数n  返回值：bool，若整数n是偶数，返回true，否则返回false  \*/  bool IsEven(int n)  {  return(n%2==0);  } |

判别素数

一个数，如果只有1和它本身两个因数，这样的数叫做[质数](https://baike.baidu.com/item/%E8%B4%A8%E6%95%B0)（或[素数](https://baike.baidu.com/item/%E7%B4%A0%E6%95%B0)）。例如 2，3，5，7 是质数，而 4，6，8，9 则不是，后者称为合成数或[合数](https://baike.baidu.com/item/%E5%90%88%E6%95%B0)。

判断某个自然数是否是素数最常用的方法就是试除法——用比该自然数的平方根小的正整数去除这个自然数，若该自然数能被整除，则说明其非素数。

C++实现功能函数：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13 | /\*函数名：IsPrime函数功能：判别自然数n是否为素数。输入值：intn，自然数n返回值：bool，若自然数n是素数，返回true，否则返回false\*/  bool IsPrime(unsignedintn)  {  unsigned maxFactor=sqrt(n);//n的最大因子  for(unsignedinti=2;i<=maxFactor;i++)  {  if(n%i==0)//n能被i整除，则说明n非素数  {  returnfalse;  }  }  return true;  } |

求最大公约数

求最大公约数最常见的方法是[欧几里德算法](https://baike.baidu.com/item/%E6%AC%A7%E5%87%A0%E9%87%8C%E5%BE%B7%E7%AE%97%E6%B3%95)（又称辗转相除法），其计算原理依赖于定理：gcd(a,b) = gcd(b,a mod b)

证明：

a可以表示成a = kb + r，则r = a mod b

假设d是a,b的一个公约数，则有d|a, d|b，而r = a - kb，因此d|r

因此d是(b,a mod b)的公约数

假设d 是(b,a mod b)的公约数，则d | b , d |r ，但是a = kb +r

因此d也是(a,b)的公约数

因此(a,b)和(b,a mod b)的公约数是一样的，其最大公约数也必然相等，得证。

C++实现功能函数：

|  |  |  |
| --- | --- | --- |
| 1  2  3  4  5  6  7 | /\*函数功能：利用欧几里德算法，采用递归方式，求两个自然数的最大公约数函数名：Gcd输入值：unsigned int a，自然数a；unsigned int b，自然数b返回值：unsigned int，两个自然数的最大公约数\*/  unsigned int Gcd(unsigned int a,unsigned int b)  {  if(b==0)  return a;  return Gcd(b,a%b);  } | |
| 1  2  3  4  5  6  7  8  9  10  11  12 | | /\*函数功能：利用欧几里德算法，采用迭代方式，求两个自然数的最大公约数函数名：Gcd输入值：unsigned int a，自然数a；unsigned int b，自然数b返回值：unsigned int，两个自然数的最大公约数\*/  unsigned int Gcd(unsigned int a,unsigned int b)  {  unsigned int temp;  while(b!=0)  {  temp=a%b;  a=b;  b=temp;  }  returna;  } | |

水仙花数

水仙花数是指一个 n 位正整数 ( n≥3 )，它的每个位上的数字的 n 次幂之和等于它本身。（例如：1^3 + 5^3+ 3^3 = 153）。

水仙花数只是[自幂数](https://baike.baidu.com/item/%E8%87%AA%E5%B9%82%E6%95%B0)的一种，严格来说三位数的3次幂数才成为水仙花数。

附：其他位数的自幂数名字

一位自幂数：独身数

两位自幂数：没有

三位自幂数：水仙花数

四位自幂数：四叶玫瑰数

五位自幂数：五角星数

六位自幂数：六合数

七位自幂数：北斗七星数

八位自幂数：八仙数

九位自幂数：九九重阳数

十位自幂数：十全十美数

假设：取1至1000内的水仙花数，那么其实只有当i>99时才成立，因为水仙花数是由3位数组成。

如果要判断一个三位数是否为水仙花数

根据运算规则，水仙花数是三位数的每个位的数的3次幂，例如999，需要取9,9,9三个数并且三数相乘的合再判断。

程序循环方式：

需要用取余数的整数的方式去完成判断条件：分别从三位数中利用取余去取百位、十位、个位数，加以判断

var a,b,c,d

for(i=1;i<1000;i++){

a = parseInt(i%10); //这一步取到了个位数

b = parseInt(i/10%10); //这一步取到了十位数

c= parseInt(i/100%10); //这一步取到了百位数

d = a\*a\*a+b\*b\*b+c\*c\*c;//水仙花数

if(d==i&&d>99){//比较判断，且是三位数。

alert(d+"是水仙花数") //输出水仙花数。

}

}

模幂运算

利用模运算的运算规则，我们可以使某些计算得到简化。

例如，我们想知道3333^5555的末位是什么。很明显不可能直接把3333^5555的结果计算出来，那样太大了。但我们想要确定的是3333^5555（%10），所以问题就简化了。

根据运算规则（4）a^b % p = ((a % p)^b) % p ，我们知道3333^5555（%10）= 3^5555（%10）。

根据运算规则（3） (a \* b) % p = (a % p \* b % p) % p ，由于5555 = 4 \* 1388 + 3，我们得到3^5555（%10）=（3^(4\*1388) \* 3^3）（%10）=（（3^(4\*1388)（%10）\* 3^3（%10））（%10）

=（（3^(4\*1388)（%10）\* 7）（%10）。

根据欧拉定理可以得到 3 ^ (4 \* k) % 10 = 1, 所以（（3^(4\*1388)（%10）\* 7）（%10）= (1 \* 7) (% 10) = 7

计算完毕。

利用这些规则我们可以有效地计算X^N(% P)。简单的算法是将result初始化为1，然后重复将result乘以X，每次[乘法](https://baike.baidu.com/item/%E4%B9%98%E6%B3%95)之后应用%运算符（这样使得result的值变小，以免溢出），执行N次相乘后，result就是我们要找的答案。

这样对于较小的N值来说，实现是合理的，但是当N的值很大时，需要计算很长时间，是不切实际的。下面的结论可以得到一种更好的算法。

如果N是偶数，那么X^N =（X\*X）^[N/2]；

如果N是奇数，那么X^N = X\*X^(N-1) = X \*（X\*X）^[N/2]；

其中[N]是指小于或等于N的最大整数。

C++实现功能函数：

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14 | /\*函数功能：利用模运算规则，采用递归方式，计算X^N(%P)函数名：PowerMod输入值：unsigned int x，底数x unsigned int n，指数nunsigned int p，模p返回值：unsigned int，X^N(%P)的结果\*/  unsignedintPowerMod(unsignedintx,unsignedintn,unsignedintp)  {  if(n==0)  {  return1;  }  unsignedinttemp=PowerMod((x\*x)%p,n/2,p);//递归计算（X\*X）^[N/2]  if((n&1)!=0)//判断n的奇偶性  {  temp=(temp\*x)%p;  }  returntemp;  } |

《孙子问题(中国剩余定理)》

在我国古代算书《[孙子算经](https://baike.baidu.com/item/%E5%AD%99%E5%AD%90%E7%AE%97%E7%BB%8F)》中有这样一个问题：

“今有物不知其数，三三数之剩二，五五数之剩三，七七数之剩二，问物几何？”意思是，“一个数除以3余2，除以5余3，除以7余2.求适合这个条件的最小数。”

这个问题称为“孙子问题”.关于孙子问题的一般解法，国际上称为“中国剩余定理”.

我国古代学者早就研究过这个问题。例如我国明朝数学家[程大位](https://baike.baidu.com/item/%E7%A8%8B%E5%A4%A7%E4%BD%8D)在他著的《算法统宗》（1593年）中就用四句很通俗的口诀暗示了此题的解法：

三人同行七十稀，五树梅花廿一枝，七子团圆正半月，除百零五便得知。

"正半月"暗指15。"除百零五"的原意是，当所得的数比105大时，就105、105地往下减，使之小于105；这相当于用105去除，求出余数。

这四句口诀暗示的意思是：当除数分别是3、5、7时，用70乘以用3除的余数，用21乘以用5除的余数，用15乘以用7除的余数，然后把这三个乘积相加。加得的结果如果比105大，就除以105，所得的余数就是满足题目要求的最小正整数解。

根据剩余定理，我把此种解法推广到有n(n为自然数）个除数对应n个余数，求最小被除数的情况。输入n个除数（除数不能互相整除）和对应的余数，计算机将输出最小被除数。

### 幂

326. Power of Three: Given an integer, write a function to determine if it is a power of three.

解：3^n%3^x=0,x<n-

// 1162261467 is 3^19, 3^20 is bigger than intreturn ( n>0 && 1162261467%n==0);

372. Super Pow：Your task is to calculate *ab* mod 1337 where *a* is a positive integer and *b* is an extremely large positive integer given in the form of an array.

## 公约数

365. Water and Jug Problem

You are given two jugs with capacities *x* and *y* litres. There is an infinite amount of water supply available. You need to determine whether it is possible to measure exactly *z* litres using these two jugs.

If *z* liters of water is measurable, you must have *z* liters of water contained within **one or both buckets** by the end.

Operations allowed:

* Fill any of the jugs completely with water.
* Empty any of the jugs.
* Pour water from one jug into another till the other jug is completely full or the first jug itself is empty.

This is a pure Math problem. We need the knowledge of number theory to cover the proof and solution. No idea why microsoft uses this problem in real interview.

**The basic idea is to use the property of Bézout's identity and check if z is a multiple of GCD(x, y)**

Quote from wiki:

Bézout's identity (also called Bézout's lemma) is a theorem in the elementary theory of numbers:

let a and b be nonzero integers and let d be their greatest common divisor. Then there exist integers x  
and y such that ax+by=d

In addition, the greatest common divisor d is the smallest positive integer that can be written as ax + by

every integer of the form ax + by is a multiple of the greatest common divisor d.

If a or b is negative this means we are emptying a jug of x or y gallons respectively.

Similarly if a or b is positive this means we are filling a jug of x or y gallons respectively.

x = 4, y = 6, z = 8.

GCD(4, 6) = 2

8 is multiple of 2

so this input is valid and we have:

-1 \* 4 + 6 \* 2 = 8

In this case, there is a solution obtained by filling the 6 gallon jug twice and emptying the 4 gallon jug once. (Solution. Fill the 6 gallon jug and empty 4 gallons to the 4 gallon jug. Empty the 4 gallon jug. Now empty the remaining two gallons from the 6 gallon jug to the 4 gallon jug. Next refill the 6 gallon jug. This gives 8 gallons in the end)

See wiki:

[Bézout's identity](https://en.wikipedia.org/wiki/B%C3%A9zout%27s_identity)

and comments in the code

**public** **boolean** **canMeasureWater**(**int** x, **int** y, **int** z) {

//limit brought by the statement that water is finallly in one or both buckets

**if**(x + y < z) **return** **false**;

//case x or y is zero

**if**( x == z || y == z || x + y == z ) **return** **true**;

//get GCD, then we can use the property of Bézout's identity

**return** z%GCD(x, y) == 0;

}

**public** **int** **GCD**(**int** a, **int** b){

**while**(b != 0 ){

**int** temp = b;

b = a%b;

a = temp;

}

**return** a;

}

## 排列组合

### Catalan

* *Cn* is the number of [Dyck words](https://en.wikipedia.org/wiki/Dyck_word)[[2]](https://en.wikipedia.org/wiki/Catalan_number#cite_note-2) of length 2*n*. A Dyck word is a [string](https://en.wikipedia.org/wiki/String_(computer_science)) consisting of *n* X's and *n* Y's such that no initial segment of the string has more Y's than X's. For example, the following are the Dyck words of length 6:

XXXYYY     XYXXYY     XYXYXY     XXYYXY     XXYXYY.

* Re-interpreting the symbol X as an open [parenthesis](https://en.wikipedia.org/wiki/Bracket#Parentheses) and Y as a close parenthesis, *Cn* counts the number of expressions containing *n* pairs of parentheses which are correctly matched:

((()))     ()(())     ()()()     (())()     (()())

* *Cn* is the number of different ways *n* + 1 factors can be completely [parenthesized](https://en.wikipedia.org/wiki/Bracket) (or the number of ways of [associating](https://en.wikipedia.org/wiki/Associativity) *n*applications of a [binary operator](https://en.wikipedia.org/wiki/Binary_operator)). For *n* = 3, for example, we have the following five different parenthesizations of four factors:

((ab)c)d     (a(bc))d     (ab)(cd)     a((bc)d)     a(b(cd))

[](https://en.wikipedia.org/wiki/File:Tamari_lattice,_trees.svg)

The [associahedron](https://en.wikipedia.org/wiki/Associahedron) of order 4 with the C4=14 full binary trees with 5 leaves

* Successive applications of a binary operator can be represented in terms of a full [binary tree](https://en.wikipedia.org/wiki/Binary_tree). (A rooted binary tree is *full* if every vertex has either two children or no children.) It follows that *Cn* is the number of full binary [trees](https://en.wikipedia.org/wiki/Tree_(graph_theory)) with *n* + 1 leaves:

[](https://en.wikipedia.org/wiki/File:Catalan_number_binary_tree_example.png)

* *Cn* is the number of non-isomorphic ordered trees with *n* vertices. (An ordered tree is a rooted tree in which the children of each vertex are given a fixed left-to-right order.)[[3]](https://en.wikipedia.org/wiki/Catalan_number#cite_note-3)
* *Cn* is the number of monotonic [lattice paths](https://en.wikipedia.org/wiki/Lattice_path) along the edges of a grid with *n* × *n* square cells, which do not pass above the diagonal. A monotonic path is one which starts in the lower left corner, finishes in the upper right corner, and consists entirely of edges pointing rightwards or upwards. Counting such paths is equivalent to counting Dyck words: X stands for "move right" and Y stands for "move up".

The following diagrams show the case *n* = 4:

[](https://en.wikipedia.org/wiki/File:Catalan_number_4x4_grid_example.svg)

This can be succinctly represented by listing the Catalan elements by column height:[[4]](https://en.wikipedia.org/wiki/Catalan_number" \l "cite_note-4)

[0,0,0,0][0,0,0,1][0,0,0,2][0,0,1,1]

[0,1,1,1] [0,0,1,2] [0,0,0,3] [0,1,1,2][0,0,2,2][0,0,1,3]

[0,0,2,3][0,1,1,3] [0,1,2,2][0,1,2,3]

[](https://en.wikipedia.org/wiki/File:Tamari_lattice,_hexagons.svg)

The triangles correspond to internal nodes of the binary trees.

* A [convex polygon](https://en.wikipedia.org/wiki/Convex_polygon) with *n* + 2 sides can be cut into [triangles](https://en.wikipedia.org/wiki/Triangle) by connecting vertices with non-crossing [line segments](https://en.wikipedia.org/wiki/Line_segment) (a form of [polygon triangulation](https://en.wikipedia.org/wiki/Polygon_triangulation)). The number of triangles formed is *n* and the number of different ways that this can be achieved is *Cn*. The following hexagons illustrate the case *n* = 4:

[](https://en.wikipedia.org/wiki/File:Catalan-Hexagons-example.svg)

* *Cn* is the number of [stack](https://en.wikipedia.org/wiki/Stack_(data_structure))-sortable [permutations](https://en.wikipedia.org/wiki/Permutation) of {1, ..., *n*}. A permutation *w* is called [stack-sortable](https://en.wikipedia.org/wiki/Stack-sortable_permutation) if *S*(*w*) = (1, ..., *n*), where *S*(*w*) is defined recursively as follows: write *w* = *unv* where *n* is the largest element in *w* and *u* and *v* are shorter sequences, and set *S*(*w*) = *S*(*u*)*S*(*v*)*n*, with *S* being the identity for one-element sequences.
* *Cn* is the number of permutations of {1, ..., *n*} that avoid the [permutation pattern](https://en.wikipedia.org/wiki/Permutation_pattern) 123 (or, alternatively, any of the other patterns of length 3); that is, the number of permutations with no three-term increasing subsequence. For *n* = 3, these permutations are 132, 213, 231, 312 and 321. For *n* = 4, they are 1432, 2143, 2413, 2431, 3142, 3214, 3241, 3412, 3421, 4132, 4213, 4231, 4312 and 4321.
* *Cn* is the number of [noncrossing partitions](https://en.wikipedia.org/wiki/Noncrossing_partition) of the set {1, ..., *n*}. [*A fortiori*](https://en.wikipedia.org/wiki/A_fortiori_argument), *Cn* never exceeds the *n*th [Bell number](https://en.wikipedia.org/wiki/Bell_number). *Cn* is also the number of noncrossing partitions of the set {1, ..., 2*n*} in which every block is of size 2. The conjunction of these two facts may be used in a proof by [mathematical induction](https://en.wikipedia.org/wiki/Mathematical_induction) that all of the *free* [cumulants](https://en.wikipedia.org/wiki/Cumulant)of degree more than 2 of the [Wigner semicircle law](https://en.wikipedia.org/wiki/Wigner_semicircle_law) are zero. This law is important in [free probability](https://en.wikipedia.org/wiki/Free_probability) theory and the theory of [random matrices](https://en.wikipedia.org/wiki/Random_matrices).
* *Cn* is the number of ways to tile a stairstep shape of height *n* with *n* rectangles. The following figure illustrates the case *n* = 4:

[](https://en.wikipedia.org/wiki/File:Catalan_stairsteps_4.svg)

* *Cn* is the number of rooted [binary trees](https://en.wikipedia.org/wiki/Binary_tree) with *n* internal nodes (*n* + 1 [leaves](https://en.wikipedia.org/wiki/Tree_(graph_theory)#Definitions) or external nodes). Illustrated in following Figure are the trees corresponding to *n* = 0,1,2 and 3. There are 1, 1, 2, and 5 respectively. Here, we consider as binary trees those in which each node has zero or two children, and the internal nodes are those that have children.

[](https://en.wikipedia.org/wiki/File:Binary_Tree.png)

* *Cn* is the number of ways to form a "mountain range" with *n* upstrokes and *n* downstrokes that all stay above a horizontal line. The mountain range interpretation is that the mountains will never go below the horizon.

[](https://en.wikipedia.org/wiki/File:Mountain_Ranges.png)

* *Cn* is the number of [standard Young tableaux](https://en.wikipedia.org/wiki/Young_tableau#Tableaux) whose diagram is a 2-by-*n* rectangle. In other words, it is the number of ways the numbers 1, 2, ..., 2*n* can be arranged in a 2-by-*n* rectangle so that each row and each column is increasing. As such, the formula can be derived as a special case of the [hook-length formula](https://en.wikipedia.org/wiki/Young_tableau#Dimension_of_a_representation).
* *Cn* is the number of ways that the vertices of a convex 2*n*-gon can be paired so that the line segments joining paired vertices do not intersect. This is precisely the condition that guarantees that the paired edges can be identified (sewn together) to form a closed surface of genus zero (a topological 2-sphere).
* *Cn* is the number of [semiorders](https://en.wikipedia.org/wiki/Semiorder) on *n* unlabeled items.[[5]](https://en.wikipedia.org/wiki/Catalan_number#cite_note-5)

## 中位数

462. Minimum Moves to Equal Array Elements II

Given a **non-empty** integer array, find the minimum number of moves required to make all array elements equal, where a move is incrementing a selected element by 1 or decrementing a selected element by 1.

算法导论：快排找中位数

## 概率

808. Soup Servings

There are two types of soup: type A and type B. Initially we have N ml of each type of soup. There are four kinds of operations:

1. Serve 100 ml of soup A and 0 ml of soup B
2. Serve 75 ml of soup A and 25 ml of soup B
3. Serve 50 ml of soup A and 50 ml of soup B
4. Serve 25 ml of soup A and 75 ml of soup B

When we serve some soup, we give it to someone and we no longer have it.  Each turn, we will choose from the four operations with equal probability 0.25. If the remaining volume of soup is not enough to complete the operation, we will serve as much as we can.  We stop once we no longer have some quantity of both types of soup.

Note that we do not have the operation where all 100 ml's of soup B are used first.

Return the probability that soup A will be empty first, plus half the probability that A and B become empty at the same time.

**Example:**

**Input:** N = 50

**Output:** 0.625

**Explanation:**

If we choose the first two operations, A will become empty first. For the third operation, A and B will become empty at the same time. For the fourth operation, B will become empty first. So the total probability of A becoming empty first plus half the probability that A and B become empty at the same time, is 0.25 \* (1 + 1 + 0.5 + 0) = 0.625.

**Notes:**

* 0 <= N <= 10^9.
* Answers within 10^-6 of the true value will be accepted as correct.

比较小的范围，用dp

无限大时，答案固定某个值

编4.1 金刚坐飞机

如果金刚坐在第n个位置，那么第i个乘客坐在自己位置的概率为f(n)

注意有个隐含条件金刚的机票座位号是1

## 拾遗

453. Minimum Moves to Equal Array Elements

Given a **non-empty** integer array of size *n*, find the minimum number of moves required to make all array elements equal, where a move is incrementing *n* - 1 elements by 1.

解：let's define sum as the sum of all the numbers, before any moves; minNum as the min number int the list; n is the length of the list;

After, say m moves, we get all the numbers as x , and we will get the following equation

sum + m \* (n - 1) = x \* n

and actually,

**x** = minNum + m

This part may be a little confusing, but @shijungg explained very well. let me explain a little again. it comes from two observations:

1. the minum number will always be minum until it reachs the final number, because every move, other numbers (besides the max) will be increamented too;
2. from above, we can get, the minum number will be incremented in every move. So, if the final number is x, it would be minNum + moves;

and finally, we will get

sum - minNum \* n = m

319. Bulb Switcher

倍数、平方数；反向思维，一个数在哪些轮会被按到

672. Bulb Switcher II

There is a room with n lights which are turned on initially and 4 buttons on the wall. After performing exactly m unknown operations towards buttons, you need to return how many different kinds of status of the n lights could be.

Suppose n lights are labeled as number [1, 2, 3 ..., n], function of these 4 buttons are given below:

1. Flip all the lights.
2. Flip lights with even numbers.
3. Flip lights with odd numbers.
4. Flip lights with (3k + 1) numbers, k = 0, 1, 2, ...

解：后面灯和前面是等价的。答案空间很小，无限循环。

754. Reach a Number

You are standing at position 0 on an infinite number line. There is a goal at position target.

On each move, you can either go left or right. During the *n*-th move (starting from 1), you take *n* steps.

Return the minimum number of steps required to reach the destination.

**Intuition**

The crux of the problem is to put + and - signs on the numbers 1, 2, 3, ..., k so that the sum is target.

When target < 0 and we made a sum of target, we could switch the signs of all the numbers so that it equals Math.abs(target). Thus, the answer for target is the same as Math.abs(target), and so without loss of generality, we can consider only target > 0.

Now let's say k is the smallest number with S = 1 + 2 + ... + k >= target. If S == target, the answer is clearly k.

If S > target, we need to change some number signs. If delta = S - target is even, then we can always find a subset of {1, 2, ..., k} equal to delta / 2 and switch the signs, so the answer is k. (This depends on T = delta / 2 being at most S.) [The proof is simple: either T <= k and we choose it, or we choose k in our subset and try to solve the same instance of the problem for T -= k and the set {1, 2, ..., k-1}.]

Otherwise, if delta is odd, we can't do it, as every sign change from positive to negative changes the sum by an even number. So let's consider a candidate answer of k+1, which changes delta by k+1. If this is odd, then delta will be even and we can have an answer of k+1. Otherwise, delta will be odd, and we will have an answer of k+2.

For concrete examples of the above four cases, consider the following:

* If target = 3, then k = 2, delta = 0 and the answer is k = 2.
* If target = 4, then k = 3, delta = 2, delta is even and the answer is k = 3.
* If target = 7, then k = 4, delta = 3, delta is odd and adding k+1 makes delta even. The answer is k+1 = 5.
* If target = 5, then k = 3, delta = 1, delta is odd and adding k+1 keeps delta odd. The answer is k+2 = 5.

**Algorithm**

Subtract ++k from target until it goes non-positive. Then k will be as described, and target will be delta as described. We can output the four cases above: if delta is even then the answer is k, if delta is odd then the answer is k+1 or k+2 depending on the parity of k.

**Complexity Analysis**

* Time Complexity: O(\sqrt{\text{target}})*O*(√​target​​​). Our while loop needs this many steps, as 1+2+⋯+k=k(k+1)21+2+⋯+k=k(k+1)2.
* Space Complexity: O(1)*O*(1).

# 系统设计

355. Design Twitter

Design a simplified version of Twitter where users can post tweets, follow/unfollow another user and is able to see the 10 most recent tweets in the user's news feed. Your design should support the following methods:

1. **postTweet(userId, tweetId)**: Compose a new tweet.
2. **getNewsFeed(userId)**: Retrieve the 10 most recent tweet ids in the user's news feed. Each item in the news feed must be posted by users who the user followed or by the user herself. Tweets must be ordered from most recent to least recent.
3. **follow(followerId, followeeId)**: Follower follows a followee.
4. **unfollow(followerId, followeeId)**: Follower unfollows a followee.

参考

**private** **static** **int** timeStamp=0;

// easy to find if user exist

**private** Map<Integer, User> userMap;

// Tweet link to next Tweet so that we can save a lot of time

// when we execute getNewsFeed(userId)

**private** **class** **Tweet**{

**public** **int** id;

**public** **int** time;

**public** Tweet next;

**public** **Tweet**(**int** id){

**this**.id = id;

time = timeStamp++;

next=null;

}

}

// OO design so User can follow, unfollow and post itself

**public** **class** **User**{

**public** **int** id;

**public** Set<Integer> followed;

**public** Tweet tweet\_head;

**public** **User**(**int** id){

**this**.id=id;

followed = **new** HashSet<>();

follow(id); // first follow itself

tweet\_head = null;

}

**public** **void** **follow**(**int** id){

followed.**add**(id);

}

**public** **void** **unfollow**(**int** id){

followed.**remove**(id);

}

// everytime user post a new tweet, add it to the head of tweet list.

**public** **void** **post**(**int** id){

Tweet t = **new** Tweet(id);

t.next=tweet\_head;

tweet\_head=t;

}

}

/\*\* Initialize your data structure here. \*/

**public** **Twitter**() {

userMap = **new** HashMap<Integer, User>();

}

/\*\* Compose a new tweet. \*/

**public** **void** **postTweet**(**int** userId, **int** tweetId) {

**if**(!userMap.containsKey(userId)){

User u = **new** User(userId);

userMap.put(userId, u);

}

userMap.**get**(userId).post(tweetId);

}

// Best part of this.

// first get all tweets lists from one user including itself and all people it followed.

// Second add all heads into a max heap. Every time we poll a tweet with

// largest time stamp from the heap, then we add its next tweet into the heap.

// So after adding all heads we only need to add 9 tweets at most into this

// heap before we get the 10 most recent tweet.

**public** List<Integer> **getNewsFeed**(**int** userId) {

List<Integer> res = **new** LinkedList<>();

**if**(!userMap.containsKey(userId)) **return** res;

Set<Integer> users = userMap.**get**(userId).followed;

PriorityQueue<Tweet> q = **new** PriorityQueue<Tweet>(users.size(), (a,b)->(b.time-a.time));

**for**(**int** user: users){

Tweet t = userMap.**get**(user).tweet\_head;

// very imporant! If we add null to the head we are screwed.

**if**(t!=null){

q.**add**(t);

}

}

**int** n=0;

**while**(!q.isEmpty() && n<10){

Tweet t = q.poll();

res.**add**(t.id);

n++;

**if**(t.next!=null)

q.**add**(t.next);

}

**return** res;

}

/\*\* Follower follows a followee. If the operation is invalid, it should be a no-op. \*/

**public** **void** **follow**(**int** followerId, **int** followeeId) {

**if**(!userMap.containsKey(followerId)){

User u = **new** User(followerId);

userMap.put(followerId, u);

}

**if**(!userMap.containsKey(followeeId)){

User u = **new** User(followeeId);

userMap.put(followeeId, u);

}

userMap.**get**(followerId).follow(followeeId);

}

/\*\* Follower unfollows a followee. If the operation is invalid, it should be a no-op. \*/

**public** **void** **unfollow**(**int** followerId, **int** followeeId) {

**if**(!userMap.containsKey(followerId) || followerId==followeeId)

**return**;

userMap.**get**(followerId).unfollow(followeeId);

}

# 未整理

凡是单调的列表，就要考虑下二分搜索。

或者存在有个有限空间。

注意结果集，或者任何集比较小的情况，

可以在此做文章。

二分、以结果为key作map

排序

快排

链表快排

插入法

两个指针法

链表归并排序

快慢指针划分

非递归归并排序

写循环注意参照for循环，不要忘了推进

高度平衡二叉树构建

公约数

裴蜀定理

https://leetcode.com/problems/water-and-jug-problem/discuss/83715/Math-solution-Java-solution

# End