# 数据结构(上)

课程不允许录像，否则将追究法律责任，赔偿损失

九章算法强化班 第2章

**扫描二维码关注微信/微博**
**获取最新面试题及权威解答**

微信: ninechapter
微博: http://www.weibo.com/ninechapter
知乎: http://zhuanlan.zhihu.com/jiuzhang
官网: http://www.jiuzhang.com

Copyright © www.jiuzhang.com

# 版权声明

九章的所有课程均受法律保护，不允许录像与传播录像
一经发现，将被追究法律责任和赔偿经济损失

Copyright © www.jiuzhang.com

# Overview

1. Union Find 并查集

2. Trie 字典树

3. Sweep Line 扫描线

Copyright © www.jiuzhang.com

# Union Find

并查集

一种用来解决集合查询合并的数据结构

支持O(1)find/O(1)union

Copyright © www.jiuzhang.com

# 并查集可以干什么？



1. 判断在不在同一个集合中。
   - find 操作

2. 关于集合合并
   - union 操作


Copyright © www.jiuzhang.com

# 并查集的操作



1. 查询 Find（递归？ 非递归？）

2. 合并 Union

Copyright © www.jiuzhang.com

# 查询 Find



- 模板代码

```
HashMap<Integer, Integer> father = new HashMap<Integer, Integer>();

int find(int x){
    int parent =  x;
    while(parent!=father.get(parent)) {
        parent = father.get(parent);
    }
    return parent;
}
```

Copyright © www.jiuzhang.com

# 合并 Union

- Key
  - 老大哥之间合并
  - 跟小弟没关系

```
HashMap<Integer, Integer> father = new HashMap<Integer, Integer>()

void union(int x, int y){
    int fa_x = find(x);
    int fa_y = find(y);
    if(fa_x != fa_y)
        father.put(fa_x, fa_y);
}
```


Copyright © www.jiuzhang.com

# 并查集完整模板



```
class UnionFind{
    UnionFind(){}
    HashMap<Integer, Integer> father = new HashMap<Integer, Integer>();

    int find(int x){
        int parent =  x;
        while(parent!=father.get(parent)) {
            parent = father.get(parent);
        }
        return parent;
    }

    void union(int x, int y){
        int fa_x = find(x);
        int fa_y = find(y);
        if(fa_x != fa_y)
            father.put(fa_x, fa_y);
    }
}
```

Copyright © www.jiuzhang.com

# Find the Connected Component in the Undirected Graph

http://www.lintcode.com/en/problem/find-the-connected-component-in-the-undirected-graph/

http://www.jiuzhang.com/solutions/find-the-connected-component-in-the-undirected-graph/

连通块：无向图一个块中节点你找得到我，我也找得到你

Copyright © www.jiuzhang.com

# ☆ Find the Connected Component in the Undirected Graph

Description | Notes | >_ Testcase | Judge

Find the number connected component in the undirected graph. Each node in the graph contains a label and a list of its neighbors. (a connected component (or just component) of an undirected graph is a subgraph in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the supergraph.)

**Notice**

Each connected component should sort by label.

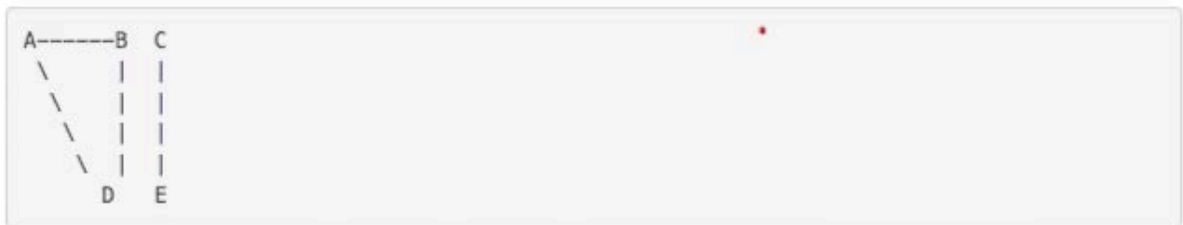
Have you met this question in a real interview? Yes

**Clarification**

Learn more about representation of graphs

**Example**

Given graph:




Return `{A,B,D}, {C,E}` . Since there are two connected component which is `{A,B,D}, {C,E}`

https://leetcode.com/problems/number-of-connected-components-in-an-undirected-graph/description/

Given `n` nodes labeled from `0` to `n - 1` and a list of undirected edges (each edge is a pair of nodes), write a function to find the number of connected components in an undirected graph.

**Example 1:**

```
     0          3
     |          |
     1 --- 2    4
```

Given `n = 5` and `edges = [[0, 1], [1, 2], [3, 4]]`, return `2`.

**Example 2:**

```
     0           4
     |           |
     1 --- 2 --- 3
```

Given `n = 5` and `edges = [[0, 1], [1, 2], [2, 3], [3, 4]]`, return `1`.

**Note:**

You can assume that no duplicate edges will appear in `edges`. Since all edges are undirected, `[0, 1]` is the same as `[1, 0]` and thus will not appear together in `edges`.

# 有向图连通块



- 弱连通块
  - 有向图一个块中，你找得到我，我可以找不到你

- 强连通块
  - 有向图一个块中，你找得到我，我也找得到你

Copyright © www.jiuzhang.com

# Find the Weak Connected Component in the Directed Graph

http://www.lintcode.com/en/problem/find-the-weak-connected-component-in-the-directed-graph/
http://www.jiuzhang.com/solutions/find-the-weak-connected-component-in-the-directed-graph/

Copyright © www.jiuzhang.com

# ☆ Find the Weak Connected Component in the Directed Graph

Description | Notes | >_ Testcase | Judge

Find the number Weak Connected Component in the directed graph. Each node in the graph contains a label and a list of its neighbors. (a connected set of a directed graph is a subgraph in which any two vertices are connected by direct edge path.)

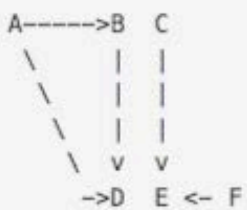
**ℹ Notice**

Sort the element in the set in increasing order

Have you met this question in a real interview? Yes

**Example**

Given graph:




Return `{A,B,D}, {C,E,F}`. Since there are two connected component which are `{A,B,D}` and `{C,E,F}`

# 路径压缩的查询 compressed_find



- 参考模板
- 平摊时间复杂度O(1)

```
int compressed_find(int x){
  int parent =  father.get(x);
  while(parent!=father.get(parent)) {
    parent = father.get(parent);
  }
  int next;

  while(x!=father.get(x)) {
    next = father.get(x);
    father.put(x, parent) ;
    x = next;
  }
  return parent;

}
```

Copyright © www.jiuzhang.com

# Google Interviewer: Number of Islands

www.lintcode.com/zh-cn/problem/number-of-islands

http://www.jiuzhang.com/solutions/number-of-islands/

Copyright © www.jiuzhang.com


Minecraft

# Google Interviewer: Number of Islands II

http://www.lintcode.com/zh-cn/problem/number-of-islands-ii/

http://www.jiuzhang.com/solutions/number-of-islands-ii/

Copyright © www.jiuzhang.com

# ☆ 岛屿的个数II

描述 笔记 数据 评测

**Given a n,m which means the row and column of the 2D matrix and an array of pair A( size k). Originally, the 2D matrix is all 0 which means there is only sea in the matrix. The list pair has k operator and each operator has two integer A[i].x, A[i].y means that you can change the grid matrix[A[i].x][A[i].y] from sea to island. Return how many island are there in the matrix after each operator.**

**注意事项**

0 代表海，1 代表岛。如果两个1相邻，那么这两个1属于同一个岛。我们只考虑上下左右为相邻。

您在真实的面试中是否遇到过这个题? Yes

样例

给定 `n = 3, m = 3`， 二元数组 A = `[(0,0),(0,1),(2,2),(2,1)]`.

返回 `[1,1,2,2]`.

标签

相关题目

# Facebook Interviewer: Graph Valid Tree

http://www.lintcode.com/problem/graph-valid-tree
http://www.jiuzhang.com/solutions/graph-valid-tree/

Union Find O(n)

Copyright © www.jiuzhang.com

Given `n` nodes labeled from `0` to `n - 1` and a list of undirected edges (each edge is a pair of nodes), write a function to check whether these edges make up a valid tree.

For example:

Given `n = 5` and `edges = [[0, 1], [0, 2], [0, 3], [1, 4]]`, return `true`.

Given `n = 5` and `edges = [[0, 1], [1, 2], [2, 3], [1, 3], [1, 4]]`, return `false`.

**Note**: you can assume that no duplicate edges will appear in `edges`. Since all edges are undirected, `[0, 1]` is the same as `[1, 0]` and thus will not appear together in `edges`.

# Surrounded Regions

http://www.lintcode.com/en/problem/surrounded-regions/

http://www.jiuzhang.com/solutions/surrounded-regions/

Copyright © www.jiuzhang.com

Given a 2D board containing `'X'` and `'O'` (the **letter** O), capture all regions surrounded by `'X'`.

A region is captured by flipping all `'O'`s into `'X'`s in that surrounded region.

For example,

```
X X X X
X O O X
X X O X
X O X X
```

After running your function, the board should be:

```
X X X X
X X X X
X X X X
X O X X
```

# 并查集考点



1、关于集合合并。

2、判断在不在同一个集合中。

Copyright © www.jiuzhang.com

# Trie

字典树

Copyright © www.jiuzhang.com

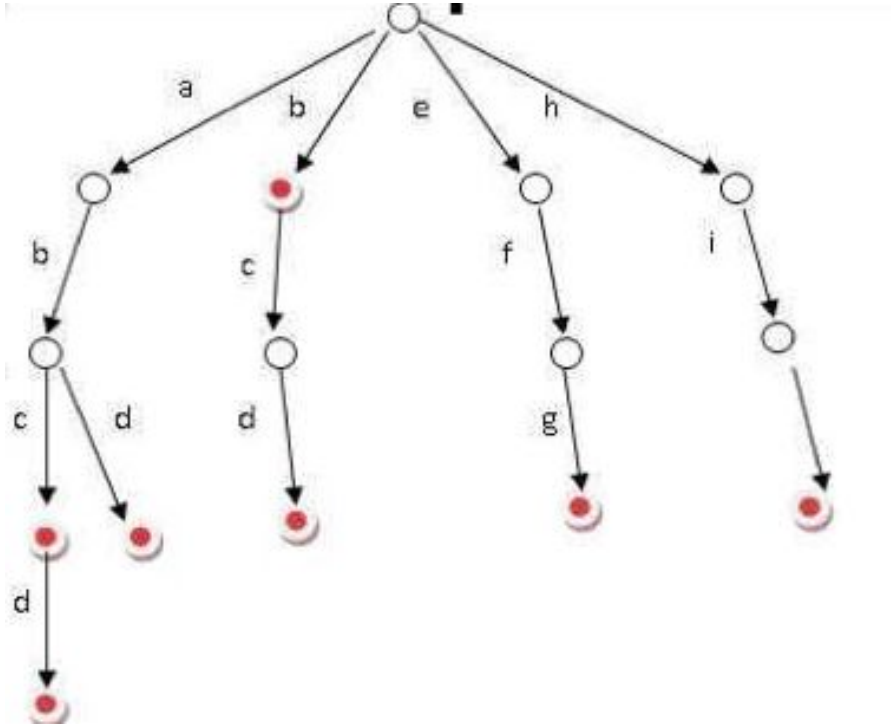
# Snapchat Interviewer: Implement Trie

http://www.lintcode.com/en/problem/implement-trie/
http://www.jiuzhang.com/solutions/trie/

Copyright © www.jiuzhang.com

# Implement Trie

- 假设有[b, abc, abd, bcd, abcd, efg, hii ]这6个单词，查找abc 在不在字典里面




Copyright © www.jiuzhang.com

# Hash vs Trie

时间复杂度Hash O(1) 是对于一个字符串

Copyright © www.jiuzhang.com

# 什么样的题目适合Trie？



- 一个一个字符串遍历

- 需要节约空间

Copyright © www.jiuzhang.com

# Microsoft Interviwer: Word Search II

http://www.lintcode.com/en/problem/word-search-ii/

http://www.jiuzhang.com/solutions/word-search-ii/

Hash vs Trie

Copyright © www.jiuzhang.com

Given a 2D board and a list of words from the dictionary, find all words in the board.

Each word must be constructed from letters of sequentially adjacent cell, where "adjacent" cells are those horizontally or vertically neighboring. The same letter cell may not be used more than once in a word.

For example,
Given **words** = `["oath","pea","eat","rain"]` and **board** =

```
[
  ['o','a','a','n'],
  ['e','t','a','e'],
  ['i','h','k','r'],
  ['i','f','l','v']
]
```

Return `["eat","oath"]`.

**Note:**

You may assume that all inputs are consist of lowercase letters `a-z`.

# Word Search II



- Given a dictionary[aca, acc] and a matrix of upper alphabets,find all words in the dictionary that can be found in the matrix.
  - acaf
  - acad
  - acae

- 解题思路:
  - 把字典建成Trie树。
  - 用dfs的方法遍历矩阵，同时在Trie上搜索前缀是否存在。
  - 查询所有Trie里面有可能出现的字符。

Copyright © www.jiuzhang.com

# Snapchat Interviewer: Add and Search Word

http://www.lintcode.com/en/problem/add-and-search-word/
http://www.jiuzhang.com/solutions/add-and-search-word/

Copyright © www.jiuzhang.com

Design a data structure that supports the following two operations:

```
void addWord(word)
bool search(word)
```

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.

For example:

```
addWord("bad")
addWord("dad")
addWord("mad")
search("pad") -> false
search("bad") -> true
search(".ad") -> true
search("b..") -> true
```

**Note:**

You may assume that all words are consist of lowercase letters `a-z`.


# Typeahead

搜索引擎

Copyright © www.jiuzhang.com

# 设计算法获得IP到城市的Map

http://www.jiuzhang.com/qa/262/

Copyright © www.jiuzhang.com

# Trie考点

- 一个一个字符串遍历
- 需要节约空间
- 查找前缀

Copyright © www.jiuzhang.com

# Sweep-Line

## 扫描线

Copyright © www.jiuzhang.com

# Amazon Interviewer: Number of Airplane in the sky

http://www.lintcode.com/en/problem/number-of-airplanes-in-the-sky/

http://www.jiuzhang.com/solutions/number-of-airplanes-in-the-sky/

Copyright © www.jiuzhang.com

# 391. Number of Airplanes in the Sky ☆

Description | Notes | >_ Testcase | Judge

Given an interval list which are flying and landing time of the flight. How many airplanes are on the sky at most?

**i** Notice

If landing and flying happens at the same time, we consider landing should happen at first.

Have you met this question in a real interview? Yes

**Example**

For interval list

```
[
  [1,10],
  [2,3],
  [5,8],
  [4,7]
]
```

Return `3`

# Number of Airplane in the sky

对于这道题有多少同学想过
按照“起点”或者“终点”对区间进行排序？

Copyright © www.jiuzhang.com

# 每日重点三题



- Number of Islands II
  - 这道题充分体现了并查集的优势
- Implement Trie
  - 理解Trie的定义和实现
- Number of Airplane in the sky
  - 扫描线入门题目

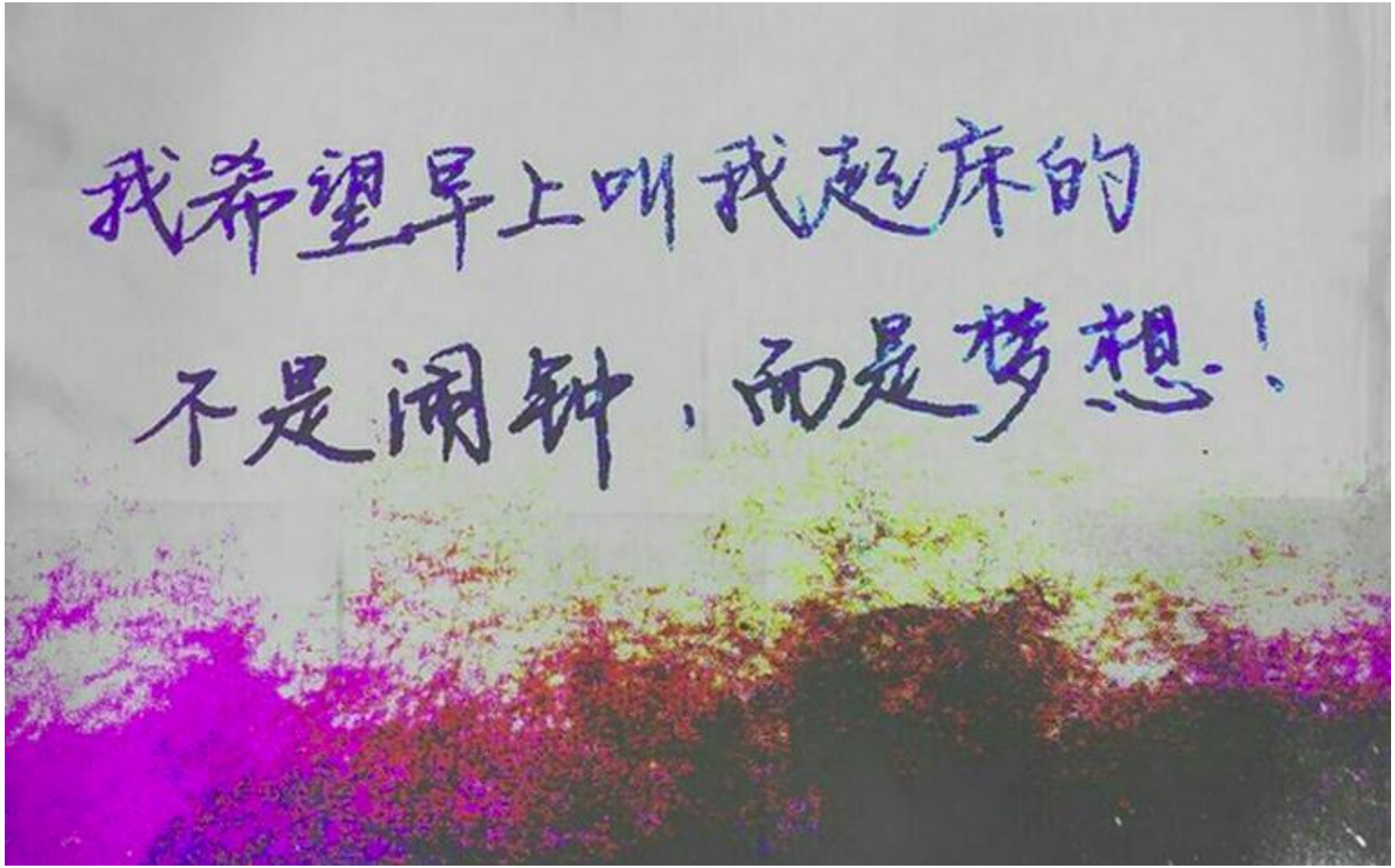
Copyright © www.jiuzhang.com

# Summary



- 数据结构的题目：
- Union Find: 集合合并，查找元素在集合里面
- Trie: 快速找到一个元素，一个字母一个字母查找
- Sweep-line: 区间拆分

Copyright © www.jiuzhang.com




Copyright © www.jiuzhang.com

# Thank You

Copyright © www.jiuzhang.com

Copyright © www.jiuzhang.com