



期末学业辅导

# 数据结构复习讲座

主讲：张嘉乐  
IEEE试点班 2018级  
[tt885.github.io](https://tt885.github.io)

2021.6



上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

# 讲座目标 | 帮助解决数据结构学习上的困难

## 困难

1. 内容抽象，没兴趣

2. 知识点太多，太杂

3. 原理懂了，应用不会

4. 书本能懂，代码薄弱



## 本讲目的

1. 介绍应用，提起兴趣

2. 梳理课程知识框架

3. 注意考试的重点

4. 深入学习方法



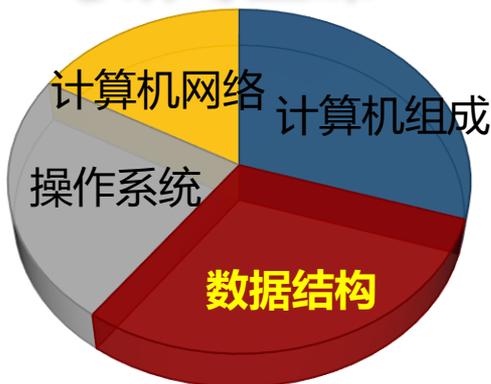
# 学习目的 | 数据结构用途——期末考完也逃不过

大学课程：

研究生面试、机考



考研专业课



其他专业课基础



课程以外：

互联网公司必考



科研基础



相关竞赛



# 故事分享 | 我是怎么学数据结构的

一年级寒假，数学建模竞赛，要用网络流模型，学会了图算法

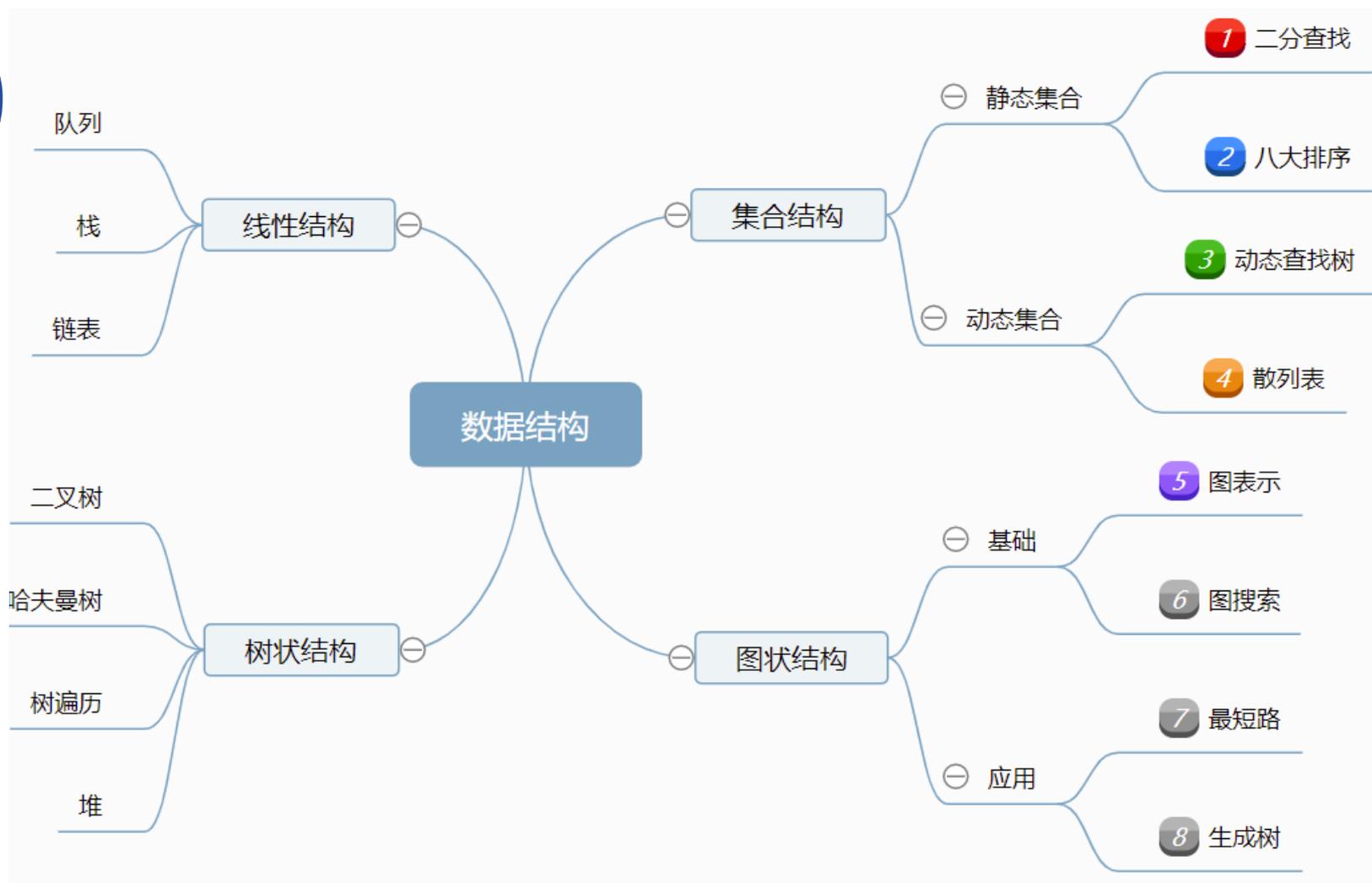
一年级春季，数据结构(俞勇)，助教布置150题，编程练几百道

二年级秋季，数据结构(张同珍)，笔试，做试卷了解不常用知识

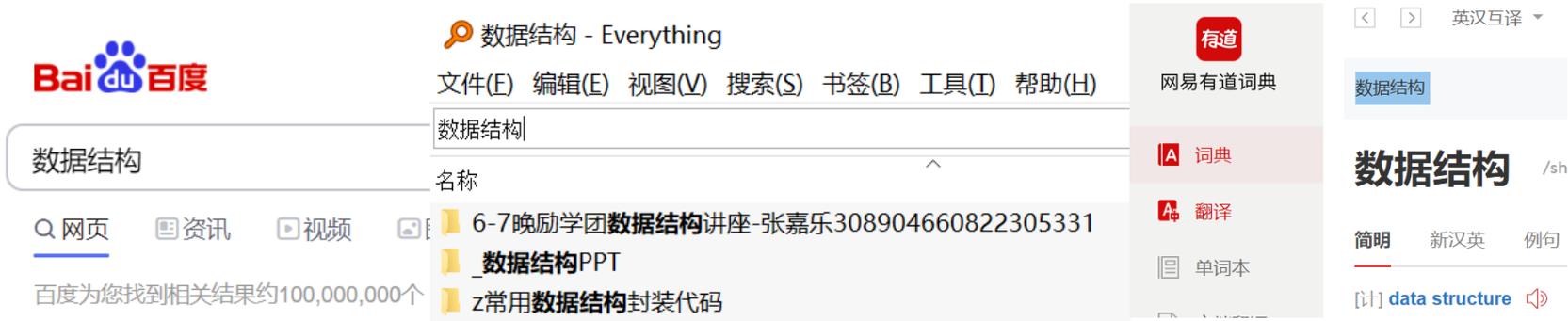
三年级开学，数学建模竞赛、培训，科研，用到各种结构和算法

三年级结束，励学团讲座，刚刚预习书本，和大家同一起跑线

# CONTENTS



# 集合结构 | 第八章：二分查找



数据结构 - Everything

文件(E) 编辑(E) 视图(V) 搜索(S) 书签(B) 工具(I) 帮助(H)

数据结构

名称

- 6-7晚励学团数据结构讲座-张嘉乐308904660822305331
- 数据结构PPT
- z常用数据结构封装代码

有道

网易有道词典

词典

翻译

单词本

数据结构 /sh

简明 新汉英 例句

data structure

## 集合，关键是查找。

求 种族值 [  ]

顺序查找: 可以

二分查找: 很好, P241

Ctrl+F: 最快, 哈希表

种族值表								
全国编号	名称	HP	攻击	防御	特攻	特防	速度	总和
370	爱心鱼	43	30	55	40	65	97	330
371	宝贝龙	45	75	60	40	30	50	300
372	甲壳龙	65	95	100	60	50	50	420
373	暴蝶螈	95	135	80	110	80	100	600
374	铁哑铃	40	55	80	35	60	30	300
375	金属怪	60	75	100	55	80	50	420
376	巨金怪	80	135	130	95	90	70	600
377	雷吉洛克	80	100	200	50	100	50	580
378	雷吉艾斯	80	50	100	100	200	50	580
379	雷吉斯奇鲁	80	75	150	75	150	50	580
380	拉帝亚斯	80	80	90	110	130	110	600
381	拉帝欧斯	80	90	80	130	110	110	600
382	盖欧卡	100	100	90	150	140	90	670
383	固拉多	100	150	140	100	90	90	670
384	烈空坐	105	150	90	150	90	95	680
385	基拉祈	100	100	100	100	100	100	600
386	代欧奇希斯 防御形态	50	70	160	70	160	90	600

# 集合结构 | 第十章：八大排序算法(内排序)

二分查找明显快，但是需要先排序。  
例：求种族值 第50的值。

分类	名称	平均复杂度	稳定性	我的偏好
插入排序	直接插入	$O(n^2)$	稳定	√在线
	希尔排序	不好分析	不稳定	
选择排序	直接选择	$O(n^2)$	不稳定	
	堆排序	$O(n \log n)$	不稳定	
交换排序	冒泡排序	$O(n^2)$	稳定	√相邻比较
	快速排序	$O(n \log n)$	不稳定	
归并排序		$O(n \log n)$	稳定	√快且稳定
基数排序		$O(d(n + r))$	稳定	

插入排序  
书P326

冒泡排序  
书P335

归并排序  
书P342

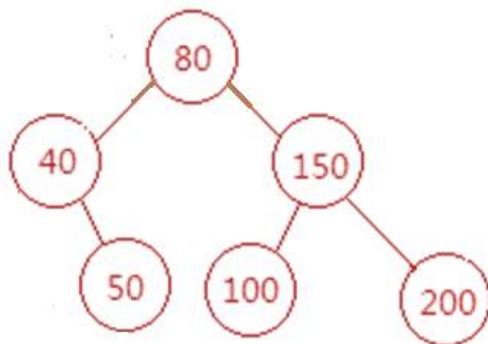
基数排序d是整数位数，r是口袋数(=10)，效率一般最高

# 集合结构 | 第九、十一章：动态查找树

数据经常需要更新——二叉查找树,书P250

目的是  $O(\log n)$ 查找,  $O(\log n)$ 插入

## 平衡树



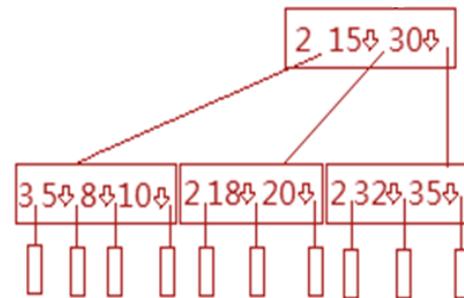
书P258

1. 插入节点45, 求之后各结点的平衡因子
2. 插入后不平衡, 求调整后的二叉平衡树

## B树

阶数  $m = 5$

书P352



1. 根和非根中间结点分别至少有几个孩子?
2. 删除关键字为18的数据后的B树。

# 集合结构 | 第九、十二章：散列表、并查集

## 散列表：O(1)查找！（理想情况）

例题:已知一个散列表大小为11，现利用散列函数  $f(\text{key})=\text{key}\%11$  将8个整数23, 19, 2, 35, 92, 66, 8, 25散列到表中，冲突时采用线性探测再散列方法。请给出散列后的散列表及其负载系数。

## 并查集：O(1)判断 “类[A]==类[B]”

优点：实现简单，比哈希表省空间

应用：最小生成树，连通图查询

全国编号	哈希值 (地址)	名称	内容
1		妙蛙种子	
2		妙蛙草	
3		妙蛙花	
4		小火龙	
5		火恐龙	
6		喷火龙	
7		杰尼龟	
8		卡咪龟	
9		水箭龟	
10		绿毛虫	
11		铁甲蛹	
12		巴大蝴	
13		独角虫	
14		铁壳昆	
15		大针蜂	
16		波波	
17		比比鸟	

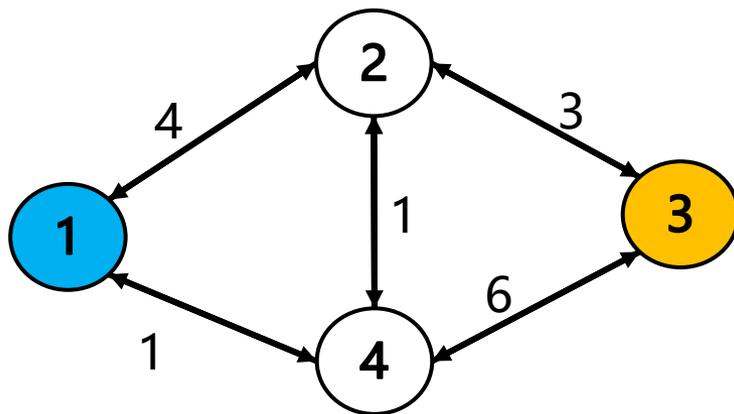
# 图状结构 | 十三章上：邻接矩阵与邻接表



图，关键问题是寻路算法。

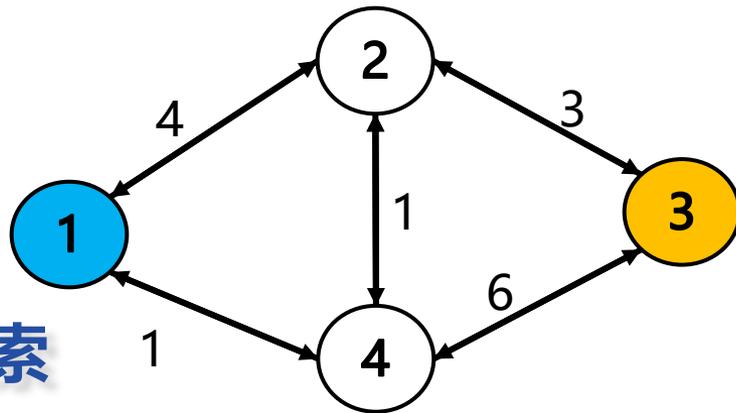
图结构怎么存到计算机里？

答：边集表示，邻接矩阵、邻接表



蓝点到黄点的最短路径长度是多少？

# 图状结构 | 十三章下：图搜索——DFS与BFS



图搜索

应用：迷宫寻路，最优解搜索

建模：需要邻接表

实现：栈(DFS)、队列(BFS)、  
优先队列(Dijkstra)

DFS



BFS



Dijkstra

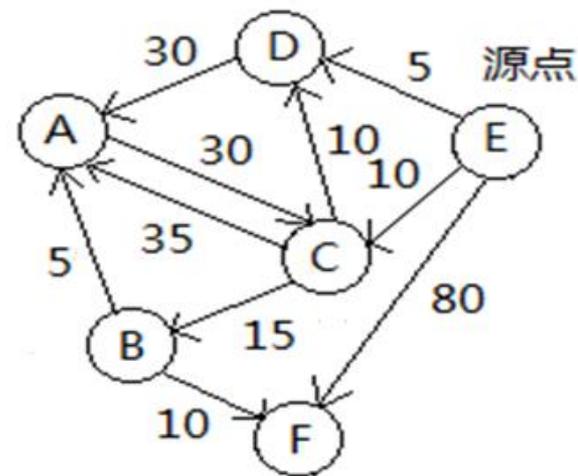
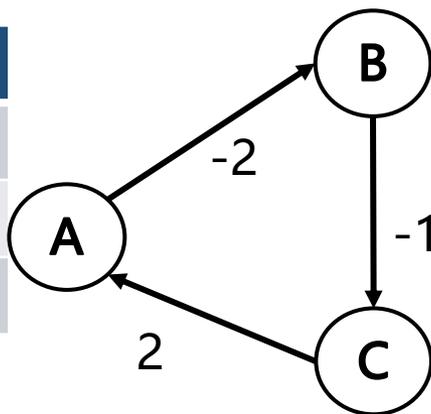


# 图状结构 | 十五章：三个最短路算法

单源最短路Dijkstra就是广度优先搜索中队列改优先队列。

Bellman和Floyd算法都是动态规划，不要求理解原理

兑换方式	汇率	lg(汇率)
A->B	1:100	-2
B->C	1:10	-1
C->A	1:0.01	2



Bellman: 书P434, 负权例子如图。

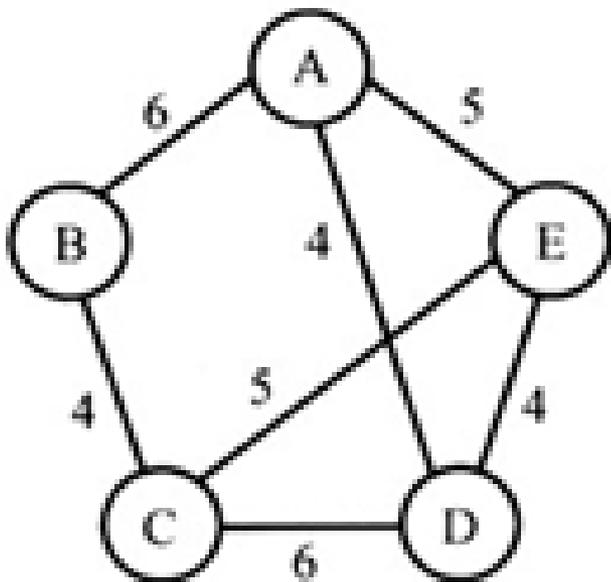
Floyd: 书P438 “迭代过程” 三重循环

求源点到各顶点最短距离，从E点到A点的最短路径。

# 图状结构 | 十四章：两个最小生成树算法

最小生成树， 假的应用——布置电线

真的应用—— 作为其他算法的组件， 如TSP



算法思想都是贪心法

Prim: 每次取最短边, 不断开

Kruskal: 每次取最短边, 不成环

# 参考文献与代码资源

[1]何进英.应用型本科院校数据结构课程教学研究与探讨[J].电脑知识与技术,2021,17(09):110-112.

[2]加州大学伯克利分校CS188人工智能课件

代码分享: <https://tt885.github.io/lectures/>

- 4个“(某数据结构).cpp”文件都是没有主函数的,只有类定义,但是类的代码都是在各种算法题中反复验证过的,应该没错。

- 另外的文件夹都是具体的算法题目,

POJ是北大系统。<http://poj.org/>

HDU是杭电系统<http://acm.hdu.edu.cn/>,后面的四位数字是题号。

SJOJ是交大系统<http://acm.sjtu.edu.cn/OnlineJudge> 不过好像坏了,ACM班自家的主页也不修修,差评。因此SJOJ题目中加入了题目说明和数据文件,直接跑代码会从data.txt中读取文件输出结果,相应的读文件代码是freopen,将控制台输入重定向为文件输入了

其他的代码也注意一下freopen这一行,去掉注释就是从文件中读取输入。交到网上评测系统时需注释这一行。

- 我认为数据结构课程考试中需要掌握到代码实现级别的算法有: 链表、二叉树遍历、归并排序、拓扑排序。也就是容易考写代码大题的知识点。

- 需要掌握到能够人工跑算法流程并输出结果,但可能不用掌握代码细节的有: 图搜索、最短路、最小生成树。容易考简答题。

并查集是一个我比较喜欢的巧妙结构,算法题中用法多,但数据结构课程很少考。