

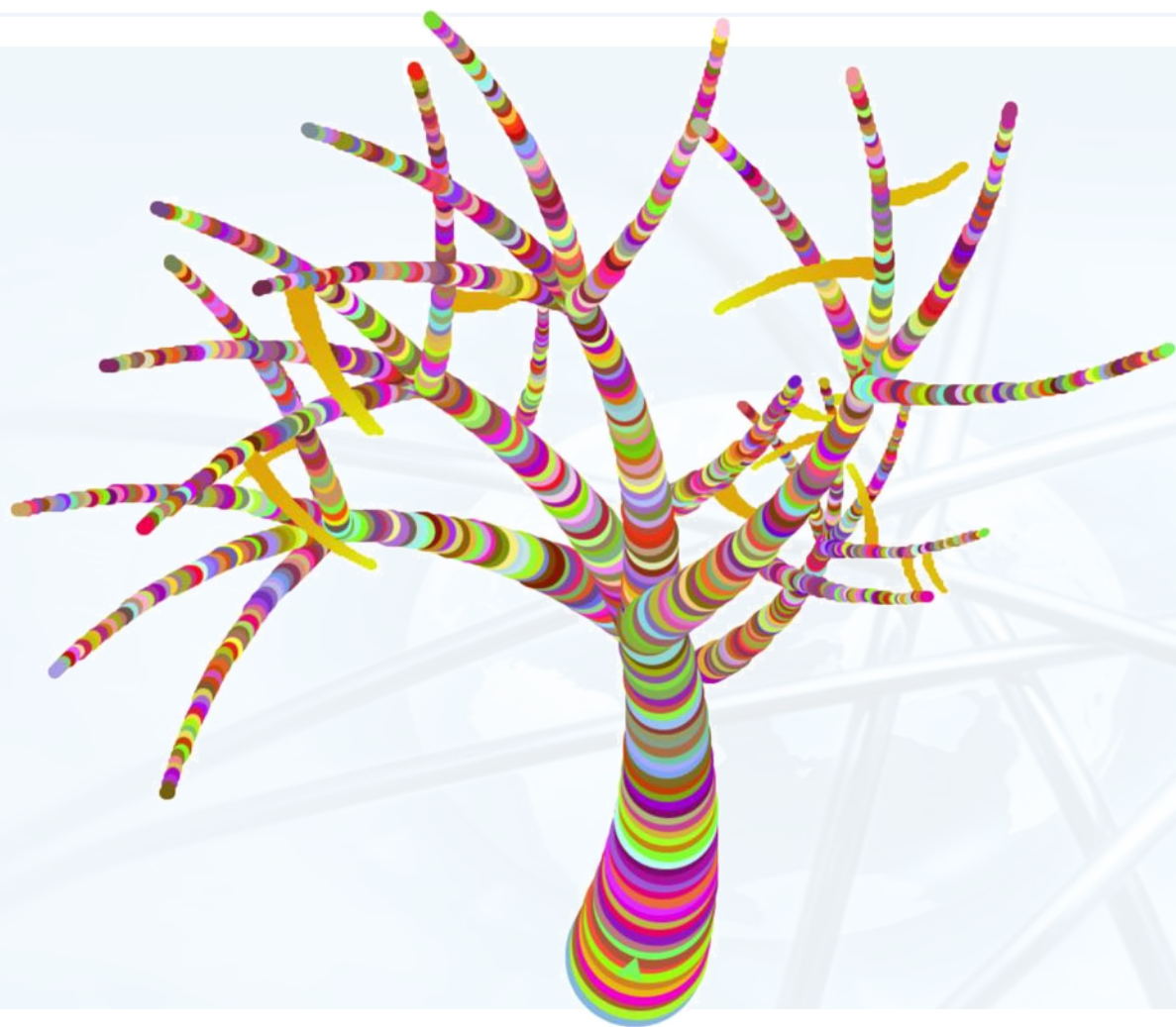


数据结构与算法（ Python ）-08/复习

陈斌 gischen@pku.edu.cn 北京大学地球与空间科学学院

目录

- ① 概述
- ② 算法分析
- ③ 基本数据结构
- ④ 递归
- ⑤ 排序与搜索
- ⑥ 树及其算法
- ⑦ 图及其算法



1概述

› 基于有穷观点的能行方法

计算的基本概念。

› 抽象计算模型-图灵机

› 计算复杂性及不可计算问题

不同问题的计算是有不同复杂度的，有些问题是无法计算的。

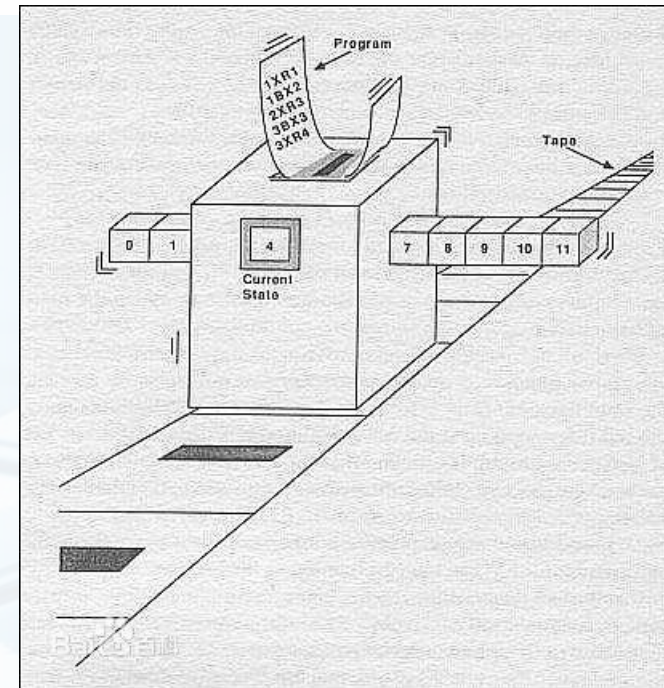
› 数据结构和抽象数据类型

抽象的概念，及什么是ADT和DS。

› 算法及衡量算法

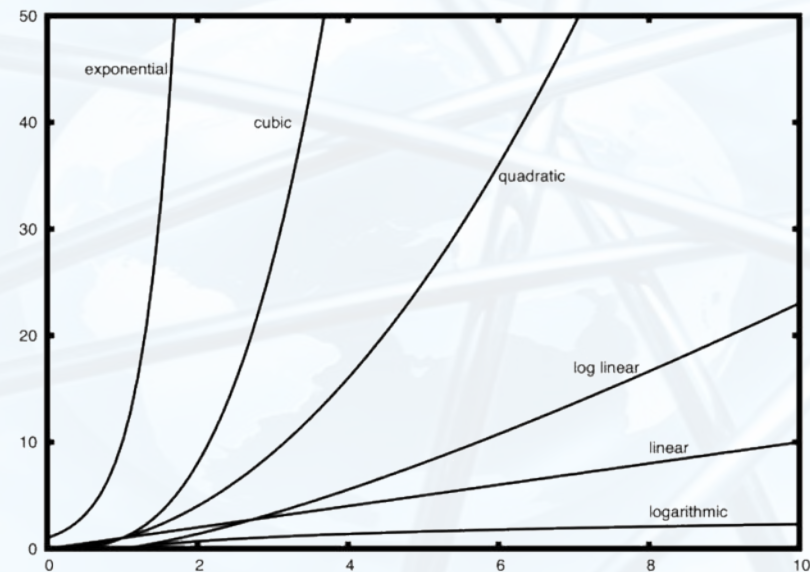
算法基本概念和衡量算法的方法。

› 基本的Python语法及实用技巧（如内置数据类型、特殊方法等）



2算法分析

- › 程序与算法的关系
- › 算法分析的概念
- › 计算资源及资源消耗指标，运行时间检测方法
- › 算法复杂度的衡量指标：大O表示法
- › 确定大O的方法及常见的大O数量级函数
 $O(1)$, $O(\log n)$, $O(n)$, $O(n \cdot \log n)$, $O(n^2)$, $O(2^n)$
- › 对算法的实现代码进行分析，以得到大O数量级
- › 理解常见Python数据类型中操作的大O数量级



3基本数据结构

› 线性数据结构的概念，理解ADT的不同实现方案及其复杂度分析

› 栈的概念、特性和ADT Stack

栈在括号匹配、进制转换及前缀中缀后缀表达式转换及求值问题的算法

› 队列的概念、特性和ADT Queue

队列在热土豆问题、打印任务问题求解的算法

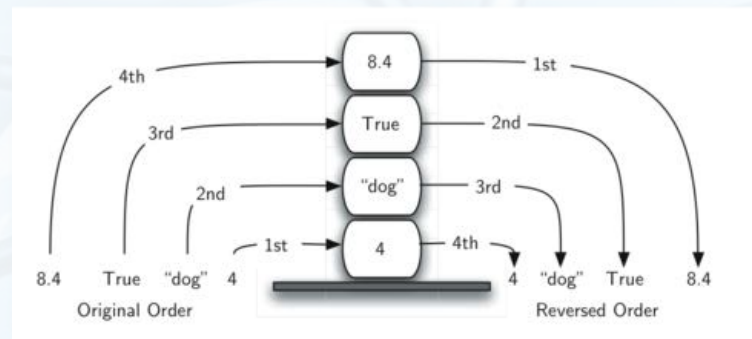
› 双端队列的概念、特性和ADT Deque

双端队列在回文词判定问题上的算法

› 列表的概念、特性和ADT List、ADT OrderedList

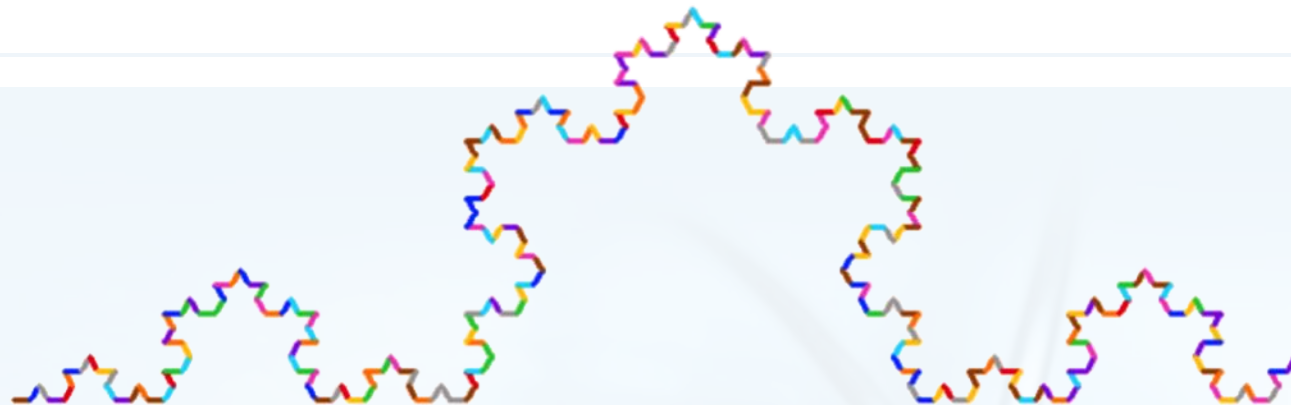
无序表的链表实现，单链表、双链表实现方案的特点

有序表的实现，及Python sort的扩展应用



4递归

- › 递归的概念及初步例子
- › 递归的“三定律”
- › 用递归解决进制转换问题
- › 递归调用的内部实现：与栈相关
- › 递归与自相似图形，理解绘制自相似图形的递归算法
- › 用递归解决河内塔问题和探索迷宫问题
- › 动态规划算法策略
- › 从兑换硬币问题对比递归算法和动态规划算法，如何避免递归爆炸



5排序与搜索

› 顺序搜索算法，以及在无序表和有序表数据结构中的不同实现

› 二分搜索算法，分而治之的算法策略

高效算法的额外开销问题，以及依据实际应用来选择算法

› 散列的概念，及散列冲突概念，完美散列函数

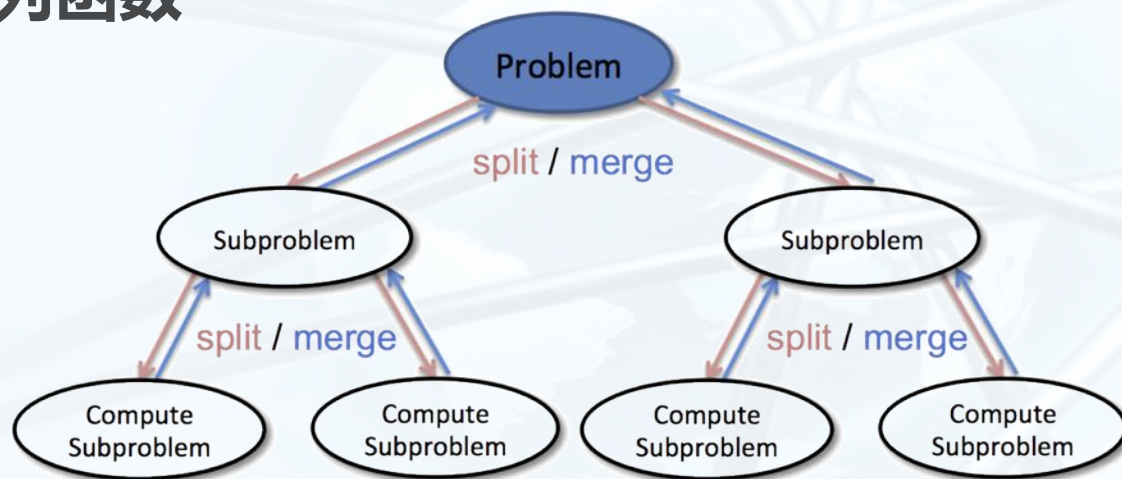
区块链技术的应用

› 散列函数设计的几种方法

› 散列冲突解决方案

开放定址法：线性探测

数据链法



› 抽象数据类型ADT Map及实现的算法分析

5排序与搜索

- › 冒泡排序算法，及性能改进
- › 选择排序算法（多趟比对，但减少交换次数）
- › 插入排序算法（为“新项”寻找插入位置，逐步扩大已排序子列表）
- › 谢尔排序算法（固定间隔的多个子列表进行插入排序，减小间隔）
- › 归并排序算法（将列表持续分裂为两半后，再合并完成排序）
- › 快速排序算法（以中值作为基准将列表分为小于和大于两部分）
- › 根据数据特征来选择排序算法

算法性能退化

如何选择一个好的算法来解决问题



6树及其算法

› 树的概念及例子，树的两种定义

熟悉树相关的术语与定义

› 实现树的方法：嵌套列表法、节点链接法

› 树的应用：解析树（语法树和表达式树）

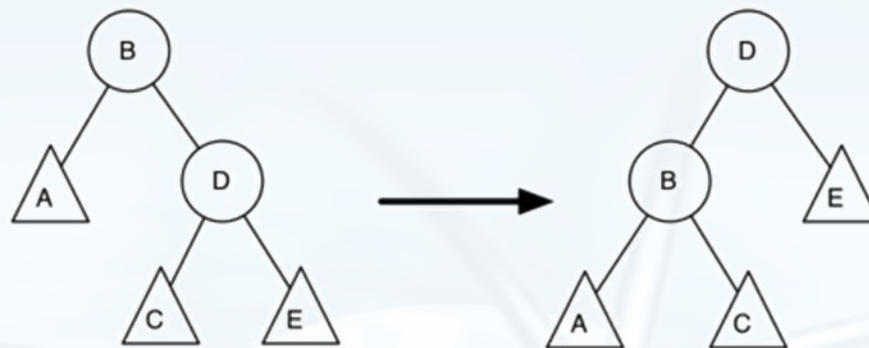
› 表达式树的建立算法，利用表达式解析树求值

› 树的遍历：前序、中序及后序遍历

在表达式生成和求值中的应用

› 优先队列的概念，实现优先队列的经典方案二叉堆

› 二叉搜索树BST及平衡树AVL树的概念及实现



7图及其算法

- › 图的概念，用图来表示的网络
- › 熟悉图的术语及定义，ADT Graph
- › 图的实现方法：邻接矩阵及邻接列表法
- › 词梯问题及广度优先搜索BFS
- › 骑士周游问题及深度优先搜索DFS
- › 通用的深度优先搜索算法
- › DFS用于解决拓扑排序和强连通分支问题
- › 路由选择，最短路径问题及Dijkstra算法
- › 信息广播，最小生成树问题及Prim算法

(附加) 人工智能经典算法

› 基因遗传算法

仿生算法

主要的流程

经典案例：拼图游戏

特点和应用

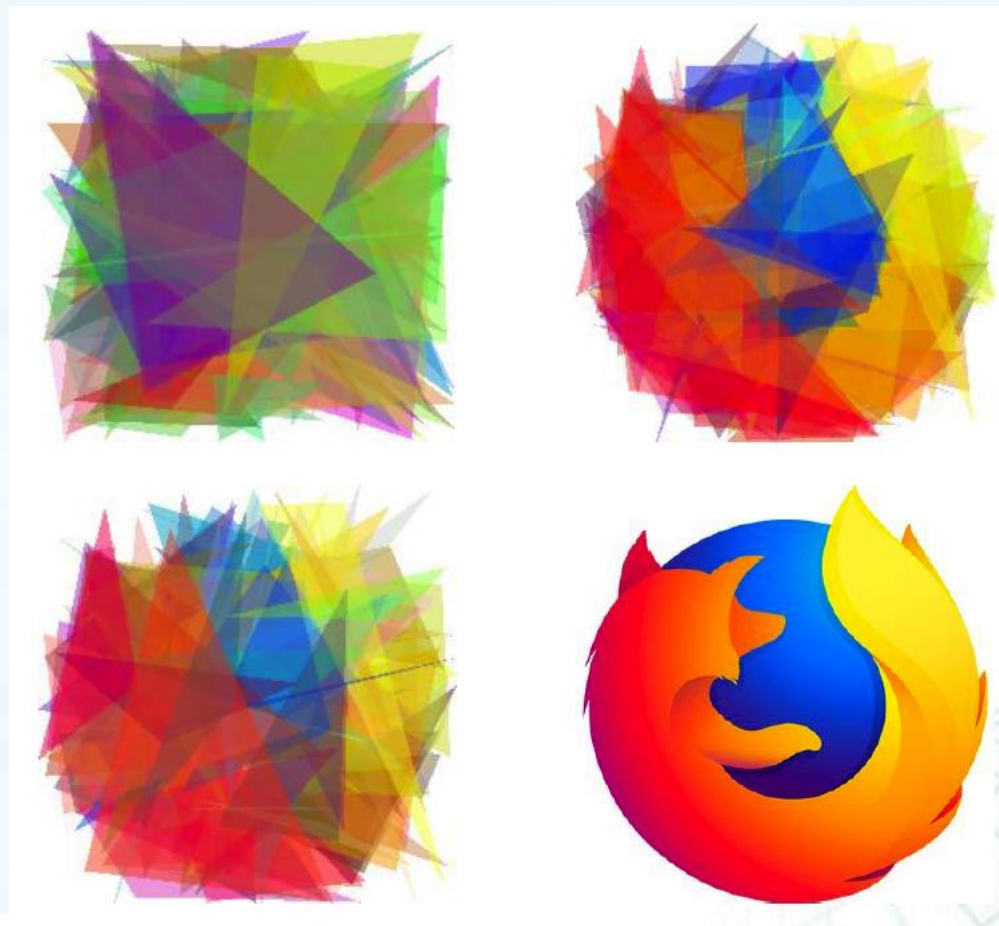
› 人工神经网络和深度学习

神经元和神经网络

主要的训练流程

经典案例：手写数字识别

特点和应用



关于期末闭卷笔试

› **时间：6月18日（周二）下午2点，二教101**

› **百分制：占总评30%**

› **三大题：判断题20分，简答题30分和综合题50分**

判断题如：有序表是一种线性数据结构

简答题如：什么是线性数据结构

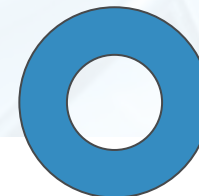
综合题如：给一个Stack，要求写算法弹出其中最大的数据项，其它数据项保持不变，算法只能用ADT Stack所提供的接口。

› **笔试中的算法用Python代码描述**

不要求语法准确运行通过，只需讲清楚算法即可。

关于考试范围

- › **能说的一点：比较侧重最后两章的内容**
 - 树及其算法
 - 图及其算法
- › **由于后两章没有DDL作业，建议将课件中的练习独立完成**



欢迎参加课程评估

› <http://kcpkg.pku.edu.cn>

