



数据结构与算法 (Python) -01/概论

刘云淮 Yunhuai.liu@pku.edu.cn

<http://www.yunhuai.net/DSA2025/CoursePage/DSA2025.html>

北京大学计算机学院

目录

- › 关于计算
计算的定义，可计算性，计算复杂度
- › 什么是计算机科学
- › 什么是数据结构
- › 什么是算法
- › 编程与算法的区别
- › 为什么学习数据结构与抽象数据类型
- › 为什么学习算法



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）

什么是算法

- › 算法是利用一个有限的指令集，遵循指令流完成特定的功能计算的狭义解释

完成从输入 $\Rightarrow$ 输出的过程

算法的基本特性：

- **有穷性**：算法经过有限步以后结束
- **确定性**：算法的下一步必须是明确的
- **可行性**：每一步都可行、且最终是正确的

不可计算问题

- › 有不少定义清晰，但无法解决的问题
并不是目前尚未找到，而是在“基于有穷观点的能行方法”的条件下，已经被证明并不存在解决方案
- › “停机问题”：判定任何一个程序在任何一个输入情况下是否能够停机
- › 不可计算数：几乎所有的无理数，都无法通过算法来确定其任意一位是什么数字

可计算数很少：如圆周率Pi，自然对数的底e

$$\pi = \frac{1}{2^6} \sum_{n=0}^{\infty} \frac{(-1)^n}{2^{10n}} \left(-\frac{2^5}{4n+1} - \frac{1}{4n+3} + \frac{2^8}{10n+1} - \frac{2^6}{10n+3} - \frac{2^2}{10n+5} - \frac{2^2}{10n+7} + \frac{1}{10n+9} \right)$$

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots$$

计算机科学研究什么

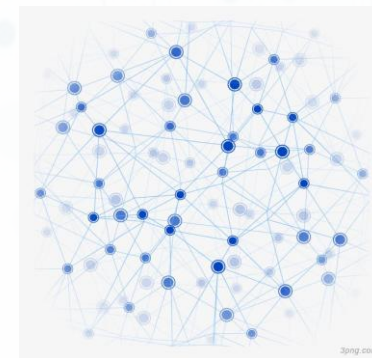
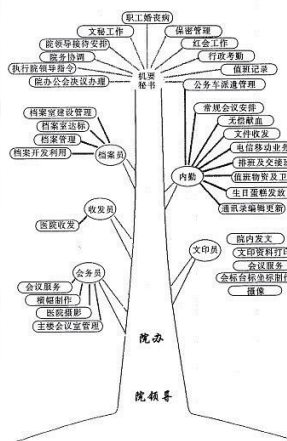
- › 计算机科学主要研究的是问题、问题解决过程，以及问题的解决方案
- › 计算复杂性理论，以及对算法的研究

数据结构的4个核心问题

- › 数据结构：按照逻辑关系组织起来的一批数据，按一定的存储结构把它存储在计算机中，并在这些数据上定义了相关运算的集合
- 数据的各种逻辑结构（关系）和存储（物理）结构，以及对应关系
 - 每种结构所适应的各种运算
 - 针对问题，设计相应的算法
 - 能分析算法的效率，包括时间和空间复杂度



- 同属色彩集合
- 
- 红色 蓝色 黄色



数据的存储结构

- › **顺序存储结构**：它是把逻辑上相邻的结点存储在物理位置相邻的存储单元里，结点间的逻辑关系由存储单元的邻接关系来体现。
- › **链式存储结构**：它不要求逻辑上相邻的结点在物理位置上亦相邻，结点间的逻辑关系是由附加的指针字段表示的。
- › **索引存储结构**：除建立存储结点信息外，还建立附加的索引表来标识结点的地址。
- › **散列存储结构**：就是根据结点的关键字直接计算出该结点的存储地址。

四种存储结构既可单独使用，又可组合使用

数据的运算

- › 定义在逻辑结构上的一系列操作以及这些操作在存储结构上的实现；
数据的运算是定义在逻辑结构上的，而具体的实现是基于存储结构。
- › 常用的运算：
检索/定位、插入、删除、修改、排序

数据结构的4个核心问题

- 数据的各种逻辑结构（关系）和存储（物理）结构，以及对应关系
- 每种结构所适应的各种运算
- 针对问题，设计相应的算法
- 能分析算法的效率，包括时间和空间复杂度



一些例题

1、 下列不属于线性结构的是：

A、 队列(queue) B、 散列表(hash table) C、 向量(vector) D、 图(graph)

2、 以下哪种结构是逻辑结构，而与存储和运算无关：

A、 队列(queue) B、 双链表(doubly linked list) C、 数组(array) D、 顺序表(Sequential list)

3、 关于算法特性描述正确的有：

A、 算法保证计算结果的正确性。

B、 组成算法的指令可以有限也可能无限。

C、 算法描述中下一步执行的步骤不确定。

D、 算法的有穷性指算法必须在有限步骤内结束。

4. 数据结构有三个基本要素:逻辑结构、存储结构以及基于结构定义的行为(运算)。下列__属于存储结构。

A:线性表 B:链表 C:字符串 D:二叉树