

# 数据结构

## 1、教材

赵敏媛，张辉、施一萍，数据结构，铁道出版社，2011年8月

## 2、主要教学内容及目的

掌握数据在计算机中的存储形式。

掌握数据的两种结构：逻辑结构和物理结构

掌握数据存储的各种物理结构：线性表、堆栈、队列、树、图

掌握数据的各种操作和算法：增加、删除、修改、读取、查找、排序

# 数据结构

## 3、考试

出勤、课堂互动、实验、作业、期末考试（期中考试）

## 4、答疑

史志才，行政楼905，时间待定

# 第一章 绪论

[1.1 数据结构的概念](#)

[1.2 算法](#)

# 本章节目录

先导课程复习

- ❖ 指针
- ❖ 结构体类型
- ❖ 链表


# 1.1 数据结构的概念

- ❖ 计算机科学的重要研究内容之一就是计算机进行数据表示和处理。这里面涉及到的问题：数据的表示、存储、传输和处理。
- ❖ 数据结构研究的主要内容是计算机所处理数据元素间的关系及其操作实现的算法，包括数据的逻辑结构、数据的存储结构以及数据的操作。

## 1.1.1基本概念和术语

- ❖ **数据(Data)**是能被计算机识别、存储和加工处理的具有一定结构的符号的总称。
- ❖ **数据项(Data Item)**是具有独立意义的不可分割的最小数据单位。 类型+数值



- 
- ❖ **数据元素**(Data Element)是数据被**使用时的基本单位**，在计算机程序中通常作为一个整体进行处理。可能包括多个数据项，如记录、节点。
  - ❖ **数据对象**(Data object)是**性质相同的数据元素的集合**，是数据的一个子集。
  - ❖ **数据结构** (Data Structure): 描述数据元素间的关系和一系列基本运算组成。

### 1.1.2逻辑结构

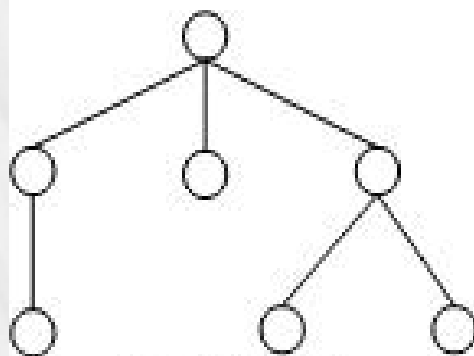
- ❖ 数据元素之间的逻辑关系称为数据的逻辑结构。
- ❖ 根据数据元素之间逻辑关系的不同特性，主要有下列**三类基本逻辑结构**。
  - ① 线性结构：结构中的数据元素之间存在一对一的关系。
  - ② 树形结构：结构中的数据元素之间存在一对多的关系。
  - ③ 图状结构：结构中的数据元素之间存在多对多的关系。

❖ 如下图所示：

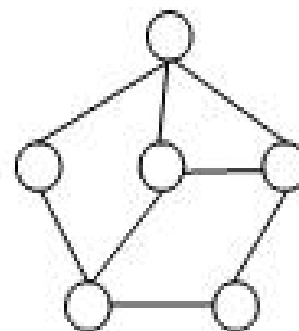
医院病人的排队数据



(a) 线性结构



(b) 树形结构



(c) 图状结构

组织结构

课程之间相互  
关系

### 1.1.3 存储结构

- ❖ 数据结构在计算机中的表示称为数据的存储结构，也称为**物理结构**。基本的存储结构有两种：**顺序存储结构**和**链式存储结构**。



- ❖ 顺序存储结构是把逻辑上相邻的数据元素存储在一组**连续的存储单元**中，其元素之间的逻辑关系由物理位置的相邻关系表示。-----C语言中的数组




- ❖ 链式存储结构将每个数据元素单独存放，称为一个结点，无需占用一整块存储空间。但为了表示结点之间的关系，需要给每个结点附加**指针字段**，用于存放下一个结点的存储地址。----C语言中的结构体与链表

### 1.1.4 抽象数据类型

- ❖ **抽象数据类型**（Abstract Data Type，简称ADT）是指一个数学模型以及定义在该模型上的一组操作。





❖ 抽象数据类型可用以下三元组表示：

$$\text{Abstract\_Data\_Type}=(D, R, P)$$

其中：

D是数据元素的有限集，

R是D上的关系的有限集，

P是对D的基本运算集。


# 1.2 算法

❖ 算法（Algorithm）是对特定问题求解步骤的一种描述。

## 1.2.1 算法的描述

❖ 三种方式：

- ① 非形式化方式：采用自然语言描述
- ② 半形式化方式：采用流程图描述
- ③ 形式化方式：采用伪代码描述



## 1.2.2 算法设计的要求

- ① 正确性。
- ② 高效率：时间短
- ③ 低存储量需求。

### 1.2.3 算法分析

- ❖ 算法的分析主要包括两个方面：算法的**时间复杂度**分析和**空间复杂度**分析。
- ❖ 一个算法是由**控制结构**和**问题的基本操作**构成的，因此，一个算法的“运行工作量”就可以用该**基本操作的重复次数**来表示。

```
void multiply(int a[n][n], int b[n][n], int c[n][n])
```

```
{ int i, j, k;
```

```
  for(i=0;i<n;i++)
```

```
    for(j=0;j<n;j++)
```

```
      { c[i][j]=0;
```

```
        for(k=0; k<n; k++)
```

```
          c[i][j]+=a[i][k]*b[k][j];
```

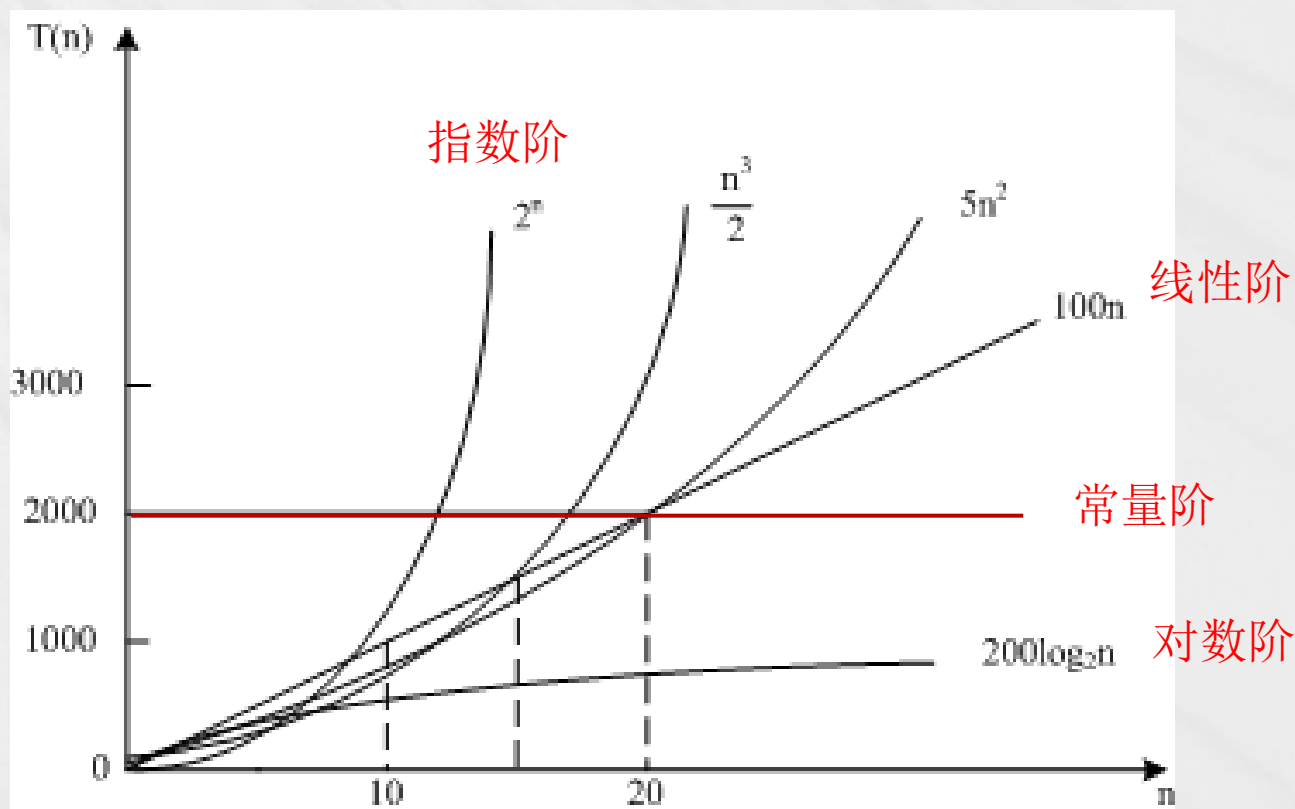
```
      } }
```

### 1.2.3 算法分析

- ❖ 一般情况下，算法中基本操作重复执行的次数是问题规模 $n$ 的某个函数，算法的时间量度记作：
  - $T(n)=O(f(n))$        $f(n)$ 为 $n$ 的代数表达式
- ❖  $T(n)$ 称做算法的渐近时间复杂度，简称时间复杂度（Time Complexity）。




- ❖ 算法的时间复杂度常见的有：
- ❖ 常量阶 $O(1)$ 、线性阶 $O(n)$ 、平方阶 $O(n^2)$ 、立方阶 $O(n^3)$ 、对数阶 $O(\log_2 n)$ 、指数阶 $O(2^n)$ 等。
- ❖ 不同数量级时间复杂度的形状如下图所示。



【例1】 **起泡排序**的算法描述如下，分析其时间复杂度。

；由小到大排序，最大的往后排□

```
❖ void bubble-sort(int a[], int n)
❖ {   int i,j,change,temp;
❖     for(i=n-1, change=1; i>=1 && change; --i)
❖     □ {   change=0;
❖           □ for(j=0;j<i;++j)
❖           □ if (a[j]>a[j+1])
❖           {   temp= a[j];  a[j] =a[j+1];  a[j+1]=temp; change=1; }
❖           □ }
❖     }//bubble-sort
```




解：在上述起泡排序算法中，问题的基本操作是“交换序列中相邻两个元素”，初始序列的状态不同，该基本操作的重复次数也有很大不同：

- ①最好情况：当初始序列为自小至大有序时，基本操作的重复次数为0，时间复杂度为 $O(1)$ ；
- ②最坏情况：当初始序列为自大至小有序时，基本操作的重复次数为：

$$1+2+3+\dots+n-1 = n(n-1)/2$$


- 时间复杂度为： $O(n^2)$

③平均情况：假设初始序列可能出现的排列情况（共 $n!$ 种）的概率相等，则时间复杂度为 $O(n^2)$ 。



通常，时间复杂度的评价指标可以分为以下三种：

- ❖ 最好时间复杂度：在最好情况下执行一个算法所需要的时间。
- ❖ 最坏时间复杂度：在最坏情况下执行一个算法所需要的时间。
- ❖ 平均时间复杂度：在平均情况下执行一个算法所需要的时间。



❖ 算法的空间复杂度（Space Complexity）是指执行算法过程中所使用的额外存储空间的开销。不包括算法程序代码和所处理的数据本身所占用的空间部分。通常，额外空间与问题的规模有关，类似于算法的时间复杂度，算法的空间复杂度记作：

□  $S(n)=O(f(n))$

□ 其中 $n$ 为问题的规模(或大小)。

由于目前计算机的内存已经足够大，一般的数据都可以处理，所以空间复杂度目前很少考虑。但是**大数据除外**！



# 本章小结

- ❖ **数据结构**是相互之间存在一种或多种特定关系的数据元素的集合。数据结构包括数据的逻辑结构、数据的存储结构以及数据的运算。
- ❖ **数据元素之间的逻辑关系**称为数据的**逻辑结构**。主要有线性结构、树形结构和图状结构。
- ❖ 数据结构在计算机中的表示称为数据的**存储结构**。基本的存储结构有**顺序存储结构**和**链式存储结构**两种。
- ❖ 算法是对特定问题求解步骤的一种描述。算法的设计既要保证正确性，同时也必须考虑算法的效率和存储量的需求。

# 本章习题

1. 什么叫数据的逻辑结构？主要有哪几种？
2. 什么叫数据的存储结构？基本的存储结构有哪几种？
3. 试述顺序存储结构和链式存储结构的区别。
4. 算法的时间复杂度和空间复杂度分别是什么？