

尚德机构

数据结构

主讲：王老师

学习是一种信仰！ IN LEARNING WE TRUST

SUNLAND

考试题型

单选题：15题 \times 2=30分

填空题：10题 \times 2=20分

解答题：4题 \times 5=20分

算法阅读题：4题 \times 5=20分

算法设计题：1题 \times 10=10分

概论



第1章

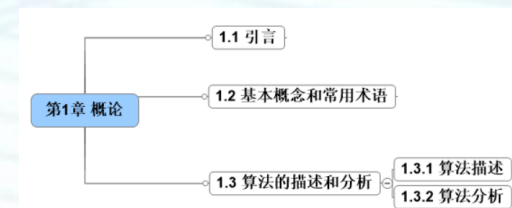
1.1 基本概念

1.2 基本概念和常用术语

1.3 算法的描述和分析

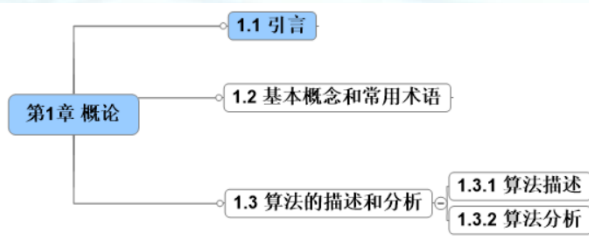
1.3.1 算法描述

1.3.2 算法分析



著名的瑞士计算机科学家沃思教授曾提出：**算法+数据结构=程序**。这里的**数据结构**指的是**数据的逻辑结构和存储结构**，而**算法**则是对数据运算的描述。由此可见，程序设计的实质是**针对实际问题选择一种好的数据结构和设计一个好的算法**，而好的算法在很大程度上取决于描述实际问题的数据结构。要设计出一个“好”的程序，就必须有好的算法，而好的算法必须建立在研究数据的特性及数据之间存在的**关系的基础之上**。

1.1第一节 引言

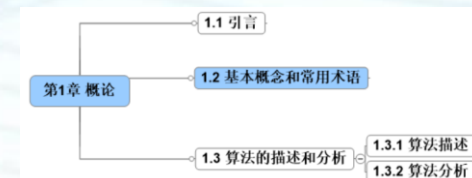


【例1.1】图书馆信息检索系统。

表1.1 图书目录卡片表

登录号	书名	作者	出版社	出版时间	分类号
01771778	数据结构	刘晓阳	高等教育	2000.08	73.961162
01509429	操作系统	许海平	机械工业	1999.12	73.752196
00592056	数据库原理	孙华英	人民邮电	2001.05	73.323265
01267435	软件工程	陈大鹏	清华大学	1998.11	73.561238
⋮	⋮	⋮	⋮	⋮	⋮

1.2 第二节 基本概念和常用术语



- **数据** (data) 是描述客观事物的数、字符以及能输入计算机中并被计算机处理的符号的集合。例如，一个代数方程的求解程序中所使用的数据是整数和实数，而一个文本编辑程序使用的数据是字符串。随着计算机的发展以及计算机应用领域的扩大，数据的含义也随之拓展了。例如，当今计算机可以处理的图形、图像、声音等也都属于数据的范畴。
- **数据元素** (data element) 是数据的基本单位。如前例中目录卡片表中的一张卡片 (**表格中的一行**)、树中的一个结点、图中的一个顶点等都是数据元素。有时一个数据元素可由若干个数据项 (**也称为字段、域、属性**) 组成，**数据项是具有独立含义的最小标识单位**，如图书卡片信息中的登录号、书名、作者等。
- **数据对象** (data object) 是具有**相同性质的数据元素的集合**，是数据的一个子集。例如，大写字母数据对象就是集合{ 'A' , 'B' , ..., 'Z' }。

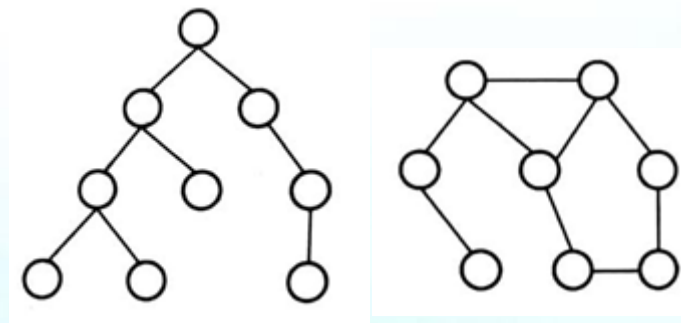
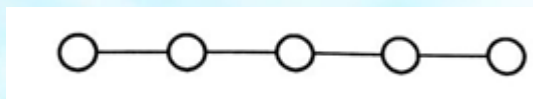
1.2 基本概念和常用术语

第1章 概论	1.1 引言	1.3.1 算法描述
	1.2 基本概念和常用术语	1.3.2 算法分析
	1.3 算法的描述和分析	

1. 数据结构包含的内容

数据结构 (data structure) 是带有结构的数据元素的集合。结构指的是数据元素之间的相互关系，即数据的组织形式，结构中的数据元素称为结点。虽然至今没有一个关于数据结构的标准定义，但它一般包括以下三个方面的内容：

- (1) 数据元素之间的**逻辑（或抽象）关系**，也称为**数据的逻辑结构**。
- **线性结构**的特征是：数据元素（结点）之间存在着**一对一**的关系，且结构中仅有一个开始结点和一个终端结点，其余结点都是仅有一个直接前趋和一个直接后继。表1.1所示就是一个典型的线性结构。本书第2章和第3章介绍的都是线性结构。
- **非线性结构**的特征是：数据元素之间存在着**一对多或多对多**的关系，即一个结点可能有多个直接前趋和多个直接后继。该结构包括**树形结构**、**图形结构**和**网状结构**等。本书的第5～7章介绍的都是非线性结构。



真题演练

下列数据结构中，逻辑结构不同的是（ ）

A: 线性表

B: 栈

C: 队列

D: 二叉树

真题演练

下列数据结构中，逻辑结构不同的是（ ）

A: 线性表

B: 栈

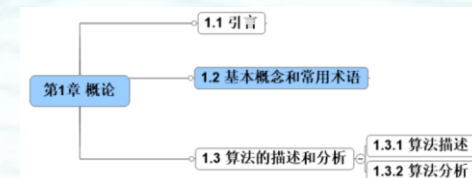
C: 队列

D: 二叉树

答案：D

线性结构的特征	数据元素（结点）之间存在着一对一的关系，且结构中仅有一个开始结点和一个终端结点，其余结点都是仅有一个直接前趋和一个直接后继。	如：线性表、栈、队列等。
非线性结构的特征	数据元素之间存在着一对多或多对多的关系，即一个结点可能有多直接前趋和多个直接后继。	如：树、图、网等。

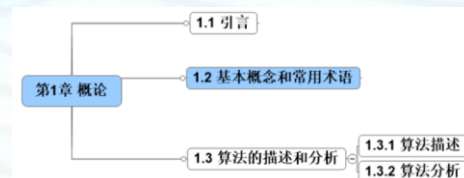
1.2 基本概念和常用术语



(2) 数据元素及其关系在计算机内的存储方式，称为数据的存储结构（物理结构）。

数据的存储结构可以用以下四种基本的存储方法实现：

①**顺序存储方法**。顺序存储方法是把逻辑上相邻的结点存储在物理位置上也相邻的连续存储单元里，由此得到的存储结构称为顺序存储结构。它通常是借助于程序设计语言的数组来描述的。该方法主要应用于线性的数据结构，但非线性的数据结构也可通过某种线性化的方法来实现顺序存储。



②**链接存储方法**。链接存储方法是用一组**不一定连续的存储单元存储逻辑上相邻的元素**，元素间的逻辑关系是由**附加的指针域**表示的，由此得到的存储结构称为链式存储结构。它通常是借助于程序设计语言中的指针来描述的。

③**索引存储方法**。索引存储方法通常是在存储元素信息的同时，还建立**附加的索引表**。表中的索引项一般形式是：（关键字，地址）。关键字是能唯一标识一个元素的一个数据项或多个数据项的组合。

④**散列存储方法**。散列存储方法的基本思想是根据**元素的关键字直接计算出该元素的存储地址**。

无论怎样定义**数据结构**，都应该将数据的**逻辑结构、存储结构及运算**这三方面看成一个整体。因此，存储结构是数据结构不可缺少的一个方面。

第1章 概论	1.1 引言
	1.2 基本概念和常用术语
	1.3 算法的描述和分析
	1.3.1 算法描述
	1.3.2 算法分析

(3) 数据的运算，即对数据元素施加的操作（行为）。

数据的运算是定义在数据的逻辑结构上的，每种逻辑结构都有一个运算的集合，最常用的运算有：检索、插入、删除、更新、排序等。数据运算是数据结构不可分割的一个方面，在给定了数据的逻辑结构和存储结构之后，按定义的运算集合及其运算性质的不同，可能导致完全不同的数据结构。例如，若对线性表的插入、删除运算限制在表的一端进行，则该线性表称为栈；若对线性表的插入运算限制在表的一端，而删除运算限制在表的另一端，则该线性表称为队列。

数据结构	数据的逻辑结构	数据元素之间的 逻辑 （或抽象） 关系 ，它与数据元素的存储结构无关，是独立于计算机的。	如：线性表、栈、队列、 树 、图、网等。
	数据的存储结构（物理结构）	数据元素及其关系在计算机内的 存储方式 。	如： 顺序 、 链式 、索引、散列等。
	数据的运算	对数据元素施加的 操作 （行为）	如：检索、插入、删除、更新、排序等。

真题演练

与数据存储结构无关的概念是 ()

A: 栈

B: 链表

C: 顺序表

D: 二叉链表

真题演练

与数据存储结构无关的概念是（）

A: 栈

B: 链表

C: 顺序表

D: 二叉链表

答案：A

数据 结构	数据的逻辑结构	数据元素之间的 逻辑 （或抽象） 关系 ，它与数据元素的存储结构无关，是独立于计算机的。	如：线性表、栈、队列、 树 、图、网等。
	数据的存储结构（物理结构）	数据元素及其关系在计算机内的 存储方式 。	如： 顺序 、 链式 、索引、散列等。
	数据的运算	对数据元素施加的 操作 （行为）	如：检索、插入、删除、更新、排序等。

真题演练

下列选项中，属于逻辑结构的是（）

A: 循环队列

B: 二叉树

C: 散列表

D: 邻接表

真题演练

下列选项中，属于逻辑结构的是（）

A: 循环队列

B: 二叉树

C: 散列表

D: 邻接表

答案：B

真题演练

下列选项中，属于逻辑结构的是（）

A: 线性表

B: 链表

C: 顺序栈

D: 循环队列

真题演练

下列选项中，属于逻辑结构的是（）

A: 线性表

B: 链表

C: 顺序栈

D: 循环队列

答案：A

真题演练

下列选项中，与数据存储结构直接相关的是（ ）

A: 线性表

B: 双向链表

C: 二叉树

D: 有向图

真题演练

下列选项中，与数据存储结构直接相关的是（ ）

A: 线性表

B: 双向链表

C: 二叉树

D: 有向图

答案：B

算法是由若干条指令组成的有穷序列，其中每条指令表示一个或多个操作。此外，算法还必须满足以下五个准则：

- (1) **输入**。算法开始前必须给算法中用到的变量初始化，一个算法的输入可以包含零个或多个数据。
- (2) **输出**。算法**至少有一个或多个输出**。
- (3) **有穷性**。算法中每一条指令的执行次数都是**有限**的，而且每一步都在有穷时间内完成，即算法必须在执行有限步后结束。
- (4) **确定性**。算法中每一条指令的含义都必须明确，无二义性。
- (5) **可行性**。算法是可行的，即算法中描述的操作都可以通过**有限次的基本运算**来实现。

真题演练

下列叙述中，不正确的是（ ）

- A: 算法解决的只能是数值计算问题
- B: 同一问题可以有多种不同算法
- C: 算法的每一步操作都必须明确无歧义
- D: 算法必须在执行有限步后结束

真题演练

下列叙述中，不正确的是（ ）

A: 算法解决的只能是数值计算问题

B: 同一问题可以有多种不同算法

C: 算法的每一步操作都必须明确无歧义

D: 算法必须在执行有限步后结束

答案：A

算法的 准则	输入	算法开始前必须给算法中用到的变量初始化，一个算法的输入可以包含零个或多个数据
	输出	算法至少有一个或多个输出
	有穷性	算法中每一条指令的执行次数都是有限的，而且每一步都在有穷时间内完成，即算法必须在执行有限步后结束
	确定性	算法中每一条指令的含义都必须明确，无二义性
	可行性	算法是可行的，即算法中描述的操作都可以通过有限次的基本运算来实现

所谓一个算法的正确性，是指对于一切合法的输入数据，该算法经过有限时间的执行都能得到正确的结果。此外，应主要考虑如下几点：

- (1) **执行算法所耗费的时间**，即**时间复杂性**。
- (2) 执行算法所耗费的存储空间，**主要是辅助空间**，即**空间复杂性**。
- (3) 算法应易于理解、易于编程，易于调试等，即可读性和可操作性。

在以上几点中，最主要的是**时间复杂性**。

【例 1.3】 百钱买百鸡问题。

公元 5 世纪末，我国古代数学家张丘建在他撰写的《算经》中提出了这样一个问题：“鸡翁一，值钱五；鸡母一，值钱三；鸡雏三，值钱一。百钱买百鸡，问鸡翁、鸡母、鸡雏各几何？”

分析：设公鸡数为 a ，母鸡数为 b ，小鸡数为 c ，根据题意可得如下的方程式：

$$\begin{cases} a+b+c=100 \\ 5a+3b+\frac{c}{3}=100 \\ c\%3=0 \end{cases}$$

算法：根据以上得出的数学模型，如果使用通常的解析法很难求解，但使用穷举法很容易实现。具体实现算法如下：

```
//函数返回值为满足问题的解数，g[]、m[]、s[]分别存放不同解的公鸡、母鸡和小鸡数
int chicken_question(int g[], int m[], int s[])
{   int a,b,c,k=0;
    for(a=0;a<=100;a++)
        for(b=0;b<=100;b++)
            for(c=0;c<=100;c++)
                if((a+b+c==100) && (5*a+3*b+c/3==100) && (c%3==0)) {
                    g[k]=a;m[k]=b;s[k]=c;
                    k=k+1;
                }
    return k;
}
```

整个算法需要执行 $101 \times 101 \times 101$

其实，对上述算法是完全可以改进的。比如，公鸡 5 元一只，百元钱全买公鸡最多也只能够买 20 只，同样全买母鸡也只能买 33 只，而小鸡只能是用买公鸡、母鸡剩余的钱来买。所以，上述算法可改为：

```
int chicken_question(int g[], int m[], int s[])
{   int a,b,c,k=0;
    for(a=0;a<=20;a++)
        for(b=0;b<=33;b++){
            c=100-a-b;
            if((5*a+3*b+c/3==100) && (c%3==0)){
                g[k]=a;m[k]=b;s[k]=c;
                k=k+1;
            }
        }
    return k;
}
```

该算法只有两重循环，内层循环体仅需要执行 $21 \times 34 = 714$ 次，与前一算法的 100 余万次相比差距相当大。因此，设计一个好的算法，对提高程序的执行效率是至关重要的。

【例1.4】求两个n阶矩阵的乘积其算法的基本操作部分如下：

```
for ( i=1; i<=n; i++)  
    for ( j=1; j<=n; j++) {  
        c[i][j] = 0;  
        for ( k=1; k<=n; k++)  
            c[i][j] = c[i][j]+a[i][k]*b[k][j];  
    }
```

$T(n)=O(n^3)$ 是矩阵乘积算法的渐近时间复杂度。

1.3.2 算法分析

1.3 算法的描述和分析

1.3.1 算法描述

1.3.2 算法分析

【例 1.5】 求下面程序段的算法时间复杂度。

```
x=0;
for (i=2;i<=n;i++)
    for (j=2;j<=i-1;j++)
        x=x+1;
```

```
x=90; y=100;
while (y>0)
    if (x>100) {
        x=x-5; y--;
    }
    else x++;
```



```
11. for (i=1;i<=n;i++) {
    y=y+1;
    for (j=1;j<=2*n;j++)
        x=x+1;
}
```

```
12. s=0;
    while (n>=(s+1)*(s+1))
        s=s+1;
```

```
13. x=1; sum=0;
    for (i=1; i<=n; i++){
        x=x*i;
        sum=sum+x;
    }
```

```
14. for(i=1;i<=n;i++)
    if(3*i<=n)
        for(j=3*i; j<=n; j++){
            x++;
            y=3*x+2;
        }
```

```
15. for(i=1;i<=n;i++)
    for(j=1;j<=i;j++)
        x=x+1;
    :
```

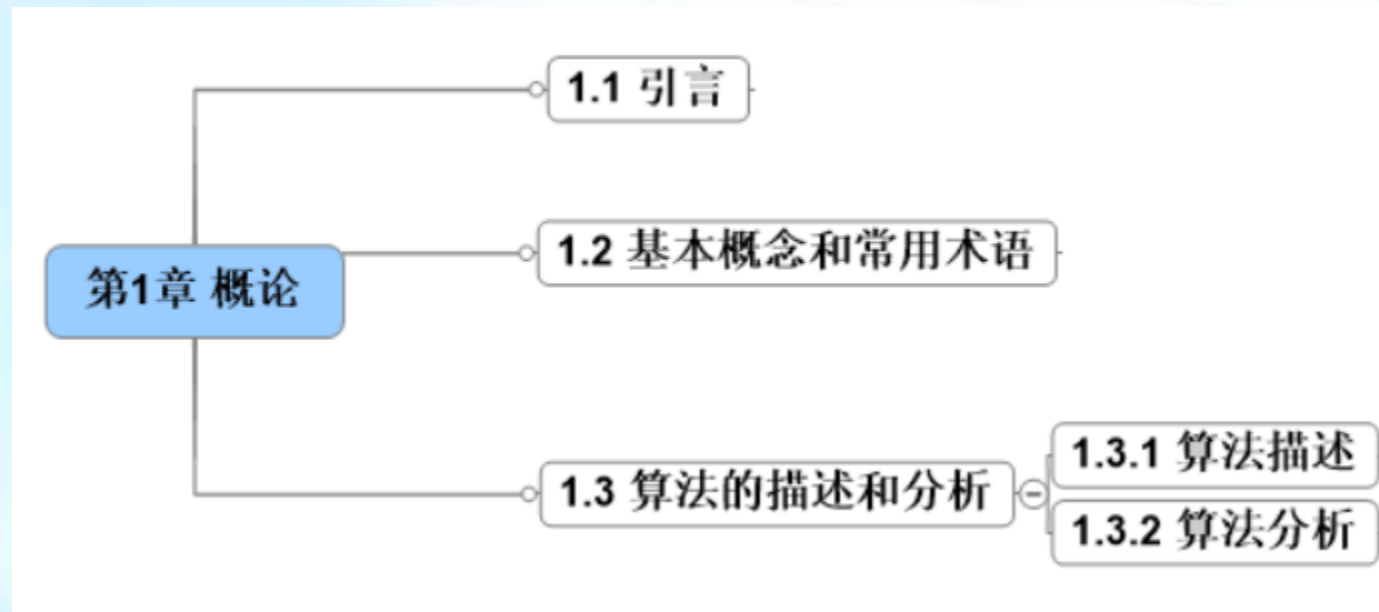
```
16. sum=0; i=0;
    while(i<=100){
        sum=sum+i;
        i++;
    }
```

```
17. x=1;s=0;
    for(i=1;i<=n;i++){
        ++x;s+=x;
    }
    for(j=1;j<=n;j++)
        for(k=1;k<=n;k++){
            x++;s=s+x;
        }
```

算法的时间复杂度通常具有 $O(1)$ 、 $O(n)$ 、 $O(\log_2 n)$ 、 $O(n \log_2 n)$ 、 $O(n^2)$ 、 $O(n^3)$ 、 $O(2^n)$ 和 $O(n!)$ 等形式，按数量级递增排列，依次为：常数阶 $O(1)$ 、对数阶 $O(\log_2 n)$ 、线性阶 $O(n)$ 、线性对数阶 $O(n \log_2 n)$ 、平方阶 $O(n^2)$ 、立方阶 $O(n^3)$ 、... k 次方阶 $O(n^k)$ 、指数阶 $O(2^n)$ 和阶乘阶 $O(n!)$ 。

类似于时间复杂度，一个算法的**空间复杂度 $S(n)$** 定义为该算法所耗费的存储空间，它是对一个算法在运行过程中**临时占用存储空间大小的度量**，是问题规模 n 的函数。一个算法在计算机存储器上所占用的存储空间，包括存储算法本身所占用的存储空间、算法的输入输出数据所占用存储空间和算法在运行过程中临时占用存储空间这三个方面。

本章总结



课后习题

一、单项选择题

1. 在数据结构中，从逻辑上可以把数据结构分为 ()
A. 紧凑结构和非紧凑结构 B. 线性结构和非线性结构
C. 内部结构和外部结构 D. 动态结构和静态结构
2. 若结点的存储地址与其关键字之间存在某种映射关系，则称这种存储结构为 ()
A. 顺序存储结构 B. 链式存储结构
C. 索引存储结构 D. 散列存储结构
3. 算法分析的两个主要方面是 ()
A. 正确性和简明性 B. 时间复杂性和空间复杂性
C. 可读性和可维护性 D. 数据复杂性和程序复杂性
4. 线性表若采用链式存储结构存储时，要求内存中可用存储单元地址 ()
A. 不一定连续的 B. 部分地址必须是连续的
C. 必须是连续的 D. 一定是不连续的

课后习题

一、单项选择题

1. 在数据结构中，从逻辑上可以把数据结构分为 ()
A. 紧凑结构和非紧凑结构 B. 线性结构和非线性结构
C. 内部结构和外部结构 D. 动态结构和静态结构
2. 若结点的存储地址与其关键字之间存在某种映射关系，则称这种存储结构为 ()
A. 顺序存储结构 B. 链式存储结构
C. 索引存储结构 D. 散列存储结构
3. 算法分析的两个主要方面是 ()
A. 正确性和简明性 B. 时间复杂性和空间复杂性
C. 可读性和可维护性 D. 数据复杂性和程序复杂性
4. 线性表若采用链式存储结构存储时，要求内存中可用存储单元地址 ()
A. 不一定连续的 B. 部分地址必须是连续的
C. 必须是连续的 D. 一定是不连续的

课后习题

5. 算法指的是

()

- A. 计算机程序
B. 解决问题的计算方法
- C. 解决问题的有限运算序列
D. 排序算法

二、填空题

6. 数据结构一般包括_____、_____和数据运算三个方面的内容。

7. 数据的逻辑结构可分为_____、_____两大类。

8. 数据的存储结构（物理结构）一般可以用_____、_____、_____及散列存储等四种存储方法表示。

9. 在选用求解一个问题的算法时,除了首先考虑算法是“正确的”之外,还主要考虑_____及算法应易于理解、易于编程、易于调试等三点。

10. 设有一批数据元素, 为了最快地存取某元素, 宜用_____结构存储, 为了方便地插入一个元素, 宜用_____结构存储。

课后习题

5. 算法指的是 ()

- A. 计算机程序
- B. 解决问题的计算方法
- C. 解决问题的有限运算序列
- D. 排序算法

二、填空题

6. 数据结构一般包括_____、_____和数据运算三个方面的内容。

7. 数据的逻辑结构可分为_____、_____两大类。

8. 数据的存储结构（物理结构）一般可以用_____、_____、_____及散列存储等四种存储方法表示。

9. 在选用求解一个问题的算法时，除了首先考虑算法是“正确的”之外，还主要考虑_____及算法应易于理解、易于编程、易于调试等三点。

10. 设有一批数据元素，为了最快地存取某元素，宜用_____结构存储，为了方便地插入一个元素，宜用_____结构存储。

一、单项选择题

1. B 2. D 3. B 4. A 5. C

二、填空题

6. 逻辑结构；存储结构

7. 线性结构；非线性结构（又可分为树形结构和图形结构）

8. 顺序存储；链式存储；索引存储

9. 执行算法所需要的时间；执行算法所需要的存储空间

10. 顺序；链式



祝大家顺利通过考试！