

《算法设计与分析》

算法设计与分析

马丙鹏

2024年09月02日



中国科学院大学

University of Chinese Academy of Sciences

课程简介

- 计算机技术专业
 - 专业硕士的核心课程
- 其他一级学科
 - 学位课或者选修课



中国科学院大学

University of Chinese Academy of Sciences

授课教师简介

■ 主讲教师

□ 马丙鹏

- 国科大，计算机科学与技术学院，教授
- 个人主页: <http://peopleucas.edu.cn/~bpma>
- 邮箱: bpma@ucas.ac.cn
- 办公地点: 学园2-387



授课教师简介

■ 助教

□ 张宏德，汪靖

➤ 国科大，计算机科学与技术学院，硕士生

➤ 邮箱：{zhanghongde22, wangjing233}
@mailsucas.ac.cn

□ 微信学习群

群聊：算法设计与分析2024 学习
研讨群



该二维码7天内(9月3日前)有效，重新进入将更新



中国科学院大学

University of Chinese Academy of Sciences 4

参考书目

■ 计算机算法设计与分析

□陈玉福，中国科学院大学讲义，2017年3月5日

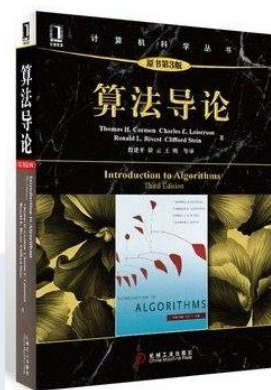
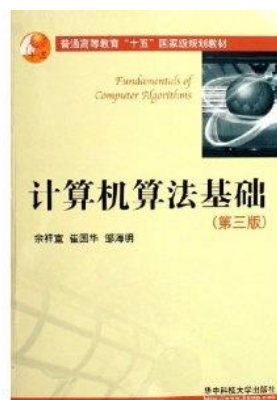
□下载地址：课程网站

■ 计算机算法基础(第三版)

□余祥宣，华中科技大学出版社，2006

■ Introduction to algorithms

□Thomas H. Cormen, etc., third edition, The MIT Press.



考核方式

■ 平时成绩

□ 作业, 20%

□ 上课, 10%

■ 期末测试

□ 闭卷笔试, 70%



中国科学院大学

University of Chinese Academy of Sciences 6

章节安排

- 第一章 算法引论
- 第二章 图与遍历算法
- 第三章 分治法
- 第四章 贪心方法
- 第五章 动态规划
- 第六章 回溯法
- 第七章 分枝—限界法
- 第八章 蛮力法
- 第九章 计算复杂性分析
- 第十章 近似算法
- 第十一章 概率算法
- 第十二章 线性规划
- 第十三章 最大流问题



序

■ 人机大战

□ 2016年3月，阿尔法围棋与围棋世界冠军、职业九段棋手李世石进行围棋人机大战，以4比1的总比分获胜。



中国科学院大学

University of Chinese Academy of Sciences

序

■ “高大上” 的算法 (Algorithm)

□ 算法是计算机软件的灵魂

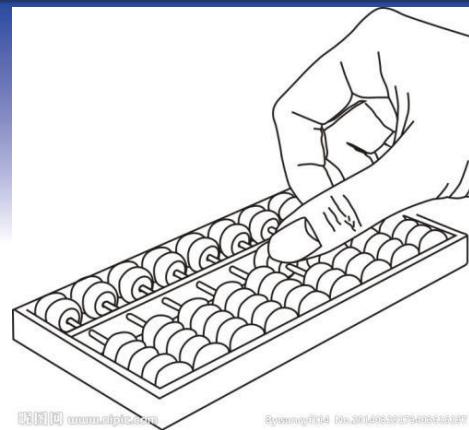
➤ 软件=数据结构+算法

➤ (Niklaus Wirth, 瑞士, 计算机科学家, 图灵奖获得者, Pascal 语言的发明者)

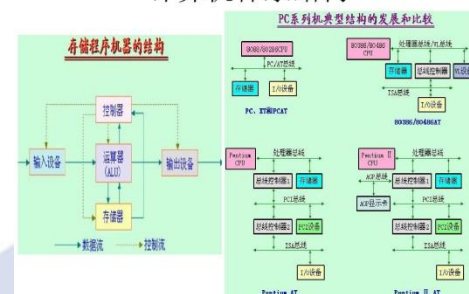
□ 算法是计算机科学的核心

➤ the core of computer science

□ 国家科技综合实力的体现



计算机体系结构



序

■ “高大上” 的算法（Algorithm）

- 20 世纪最伟大的科学技术发明---计算机；
- 计算机是对人脑的模拟，它强化了人的思维；
- 没有软件的支持，超级计算机只是一堆废铁而已。



序

■ 现代科学研究的三大支柱



中国科学院大学

University of Chinese Academy of Sciences 11

序

■ 算法不陌生

□ 举个例子：排序（Sort）

- 问题描述：将一组数按照从小到大的顺序整理有序
- 基本思想：小的数往前排，大的数往后排

□ 怎么排？

- 冒泡排序
- 选择排序
- 归并排序
- 快速排序
- 堆排序
-

- 每种算法都是一组特定的规则，规定了一种处理数据的运算方法
- 数据是死的，算法是活的



序

■ 生活中的算法

衣



食



美团

饿了么



大众点评
dianping.com

住



行



中国科学院大学

University of Chinese Academy of Sciences

13 清华大学出版

序

■ 算法不陌生

□ $1 + 2 + 3 + 4 + \dots + 100 = ?$

□ 方法一：

➤ $1 + 2 = 3$

➤ $3 + 3 = 6$

➤ $6 + 4 = 10$

➤ ...

➤ $4950 + 100 = 5050$

□ 方法二：

➤ $(1 + 100) * 50 = 5050$

显然方法二更高效！



中国科学院大学

University of Chinese Academy of Sciences 14

序

■ 算法不陌生

□ 既然每种算法都可以实现排序的目的，何必费心研究这么多排序算法？

- ① 好奇心的需要：人们热衷于寻找不同的方法解决各种各样的问题。
- ② 研究的需要：算法和算法是不同的。
 - 性质不同，稳定不稳定
 - 性能不同，速度、空间
 - 适用的场合不同
- ③ 应用的需要：没有万能的算法适用于所有的应用。

怎么选择？根据性能，结合需求，综合来选。

如何了解每种算法的性能？算法分析



算法的由来



公元前300年
辗转相除法
《几何原本》

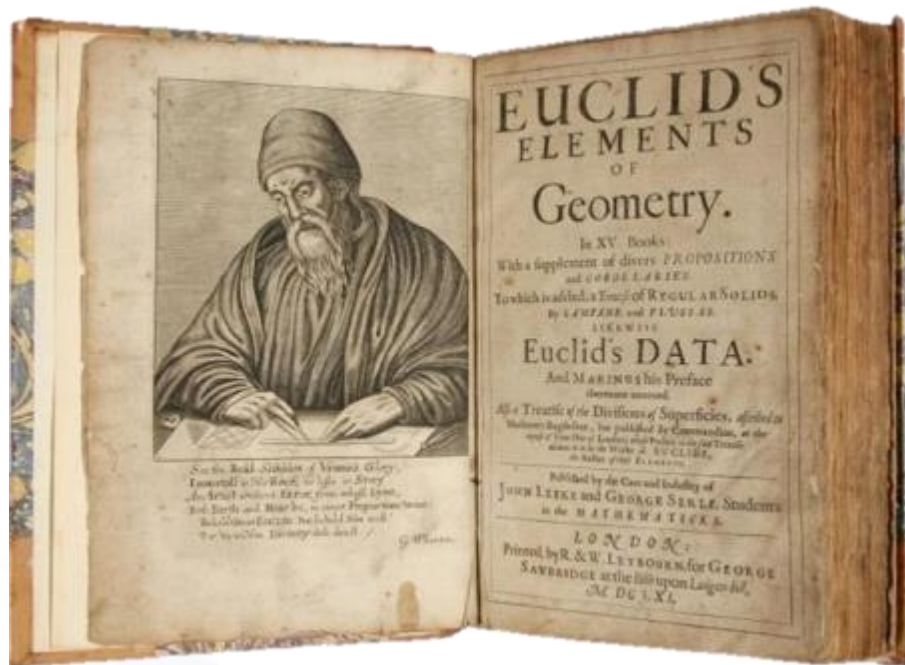


算法的由来

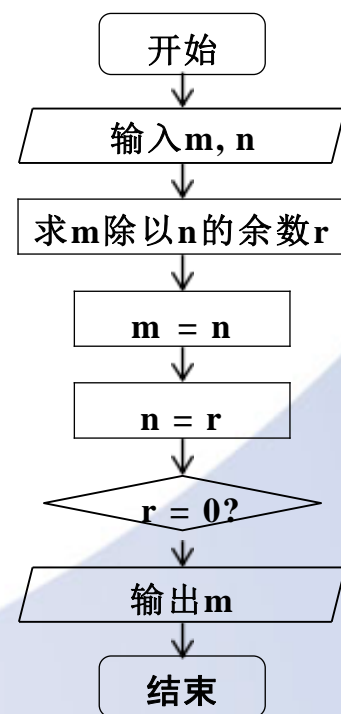
■ 辗转相除法

□ 用于计算两个整数的最大公约数

□ 约公元前300年由欧几里德提出



欧几里得《几何原本》



辗转相除法流程图



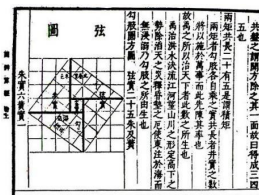
中国科学院大学

University of Chinese Academy of Sciences 17

算法的由来



公元前300年
辗转相除法
《几何原本》



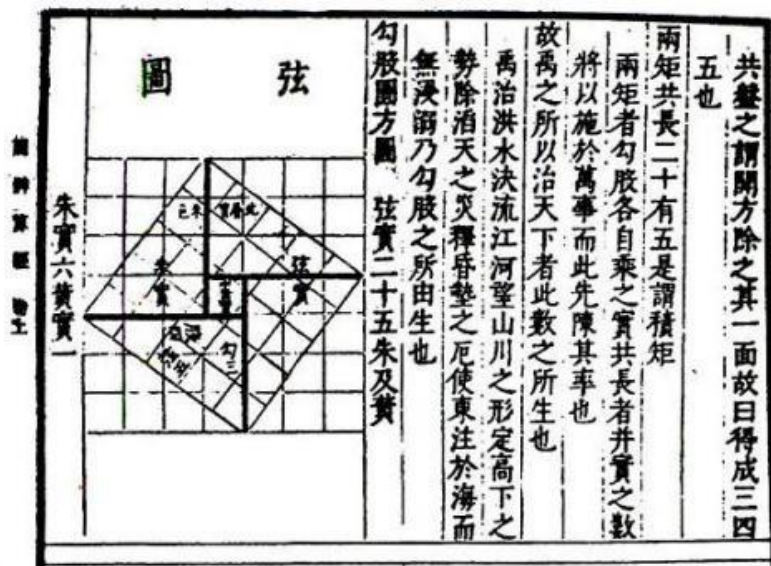
公元前1世纪
算法
《周髀算经》



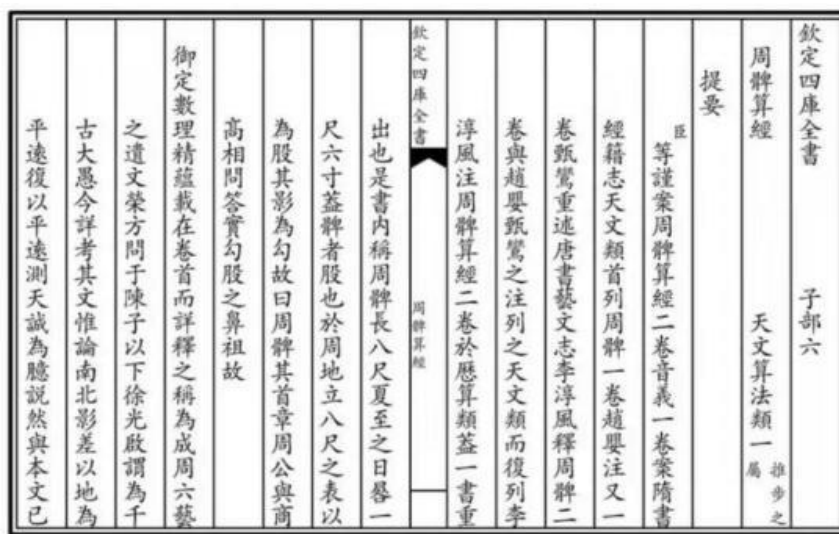
算法的由来

■ 名称由来

□ 中文名称“算法”出自约成书于公元前1世纪的《周髀算经》



介绍了勾股定理：“以日下为勾，日高为股，勾股各自乘，并而开方除之”



采用简便可行方法确定天文历法
揭示日月星辰运行规律



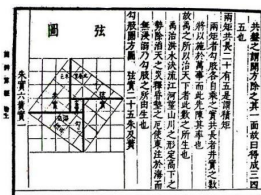
中国科学院大学

University of Chinese Academy of Sciences 19

算法的由来



公元前300年
辗转相除法
《几何原本》



公元前1世纪
算法
《周髀算经》



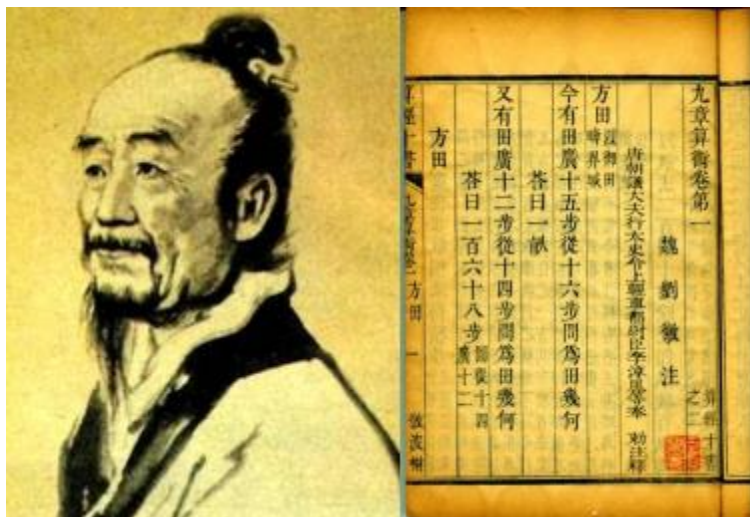
魏晋时期
割圆术
《九章算术注》



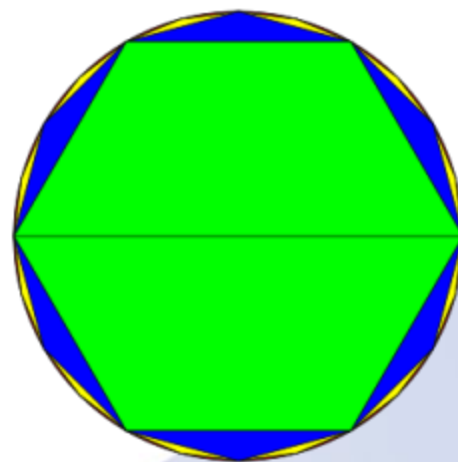
算法的由来

■ 割圆术

- 内接正多边形去无限逼近圆，以此求取圆周率的方法
- 魏晋时期的数学家刘徽在《九章算术注》中首创



刘徽与《九章算术注》



割圆术示意图



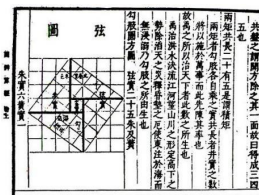
中国科学院大学

University of Chinese Academy of Sciences 21

算法的由来



公元前300年
辗转相除法
《几何原本》



公元前1世纪
算法
《周髀算经》



魏晋时期
割圆术
《九章算术注》



7世纪
《算经十书》



算法的由来

■ 算经十书

□ 公元656年唐高宗规定将十部汉、唐一千多年间的十部著名数学著作作为国家最高学府的算学教科书，用以进行数学教育和考试，后世通称为《算经十书》



唐高宗李治



周髀算经



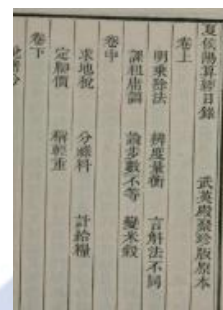
九章算术



海岛算经



张丘建算经



夏侯阳算经



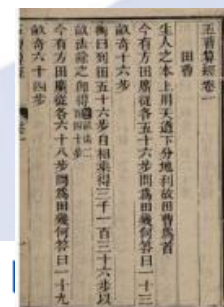
五经算术



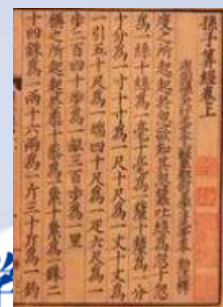
缉古算经



缀术



五曹算经



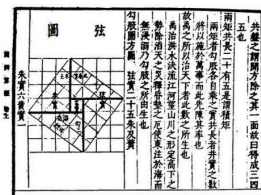
孙子算经



算法的由来



公元前300年
辗转相除法
《几何原本》



公元前1世纪
算法
《周髀算经》



魏晋时期
割圆术
《九章算术注》



7世纪
《算经十书》



9世纪
Algorithm
阿尔·花拉子密



中国科学院大学
University of Chinese Academy of Sciences

算法的由来

■ 名称由来

□ 波斯著名的数学家、天文学家、地理学家阿尔·花拉子密在公元825年写成《印度数字算术》一书，对于印度-阿拉伯数字系统在中东及欧洲的传播起到了重要作用

□ 该书被翻译成拉丁语
“*Algoritmi de numero Indorum*”，
花拉子密的拉丁文音译即为“算法”（Algorithm）一词的由来



苏联在1983年发行邮票纪念花拉子密1200岁生辰



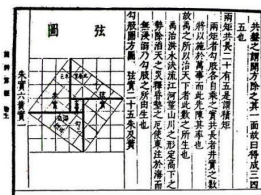
中国科学院大学

University of Chinese Academy of Sciences 25

算法的由来



公元前300年
辗转相除法
《几何原本》



公元前1世纪
算法
《周髀算经》



魏晋时期
割圆术
《九章算术注》



7世纪
《算经十书》



9世纪
Algorithm
阿尔·花拉子密



20世纪30年代~40年代
艾伦图灵与冯诺依曼
中国科学院大学
University of Chinese Academy of Sciences

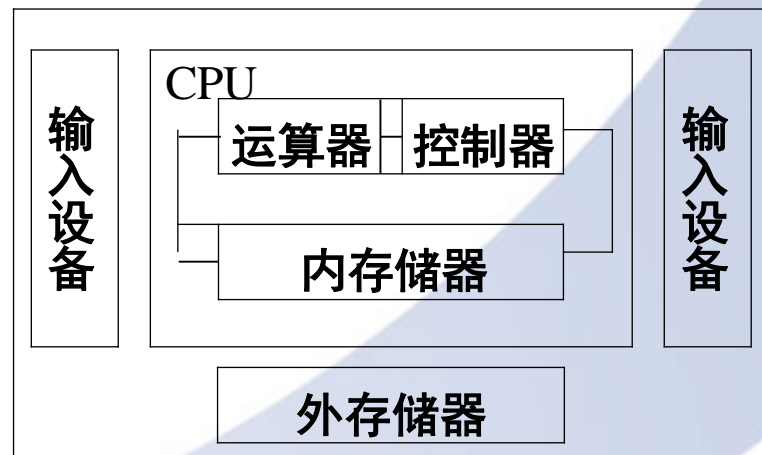
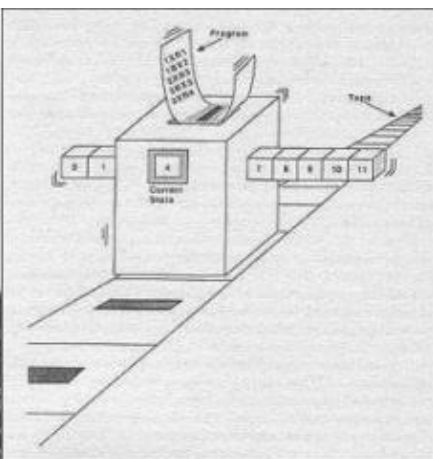


算法的由来

■ 算法与计算机的结合

□1936年，艾伦·图灵提出图灵机，通过建立通用计算机模型，刻画计算机的计算行为

□1946年，冯·诺依曼提出存储程序原理



理论计算机科学与人工智能之父
艾伦·图灵
Alan Turing

现代计算机之父
约翰·冯·诺伊曼
John von Neumann
中国科学院大学
University of Chinese Academy of Sciences 27

序

■ 学习算法你能获得什么

- 积累丰富的解决问题的经验和技能

- 训练编程

 - 从简单技术的应用到技巧、技能的提高

- 训练思维

 - 严谨、科学的逻辑推理能力

 - 培养计算机系统观，培养计算机思维

 - 培养使用计算机问题求解能力



序

■ 课程核心：

- 介绍算法设计与分析的基本理论、方法和技术，训练程序思维，奠定算法设计的基础。

■ 教学目的：

- 在理论学习上，掌握算法设计与分析的基本理论和方法，培养设计算法和分析算法的能力。
- 在实践教学上，掌握算法实现的技术、技巧，学习算法的正确性验证、效率分析、优化技术，以及算法在实际问题中的分析与应用。
- 培养独立研究和创新能力。



序

■ 本课程需要的基础

- 数据结构 ✓
- 程序设计语言（C/C++）：结构化设计 ✓
- 一定的数学基础：高数、离散、概率 ✓
- 一些背景知识：操作系统、编译



序

■ 与其他课程的关系

数据结构

算法设计与分析

高级程序设计语言 (C, C++)

与数据结构的区别:

●考虑问题的**角度**: 数据结构关心不同的数据结构在解题中的作用和效率; 算法关心不同设计技术的适用性和效率。

●考虑问题的**高度**: 数据结构关心的是解具体问题, 算法不仅如此, 它提供一种解决问题的通用方法。



中国科学院大学

University of Chinese Academy of Sciences 31

第一章 算法引论

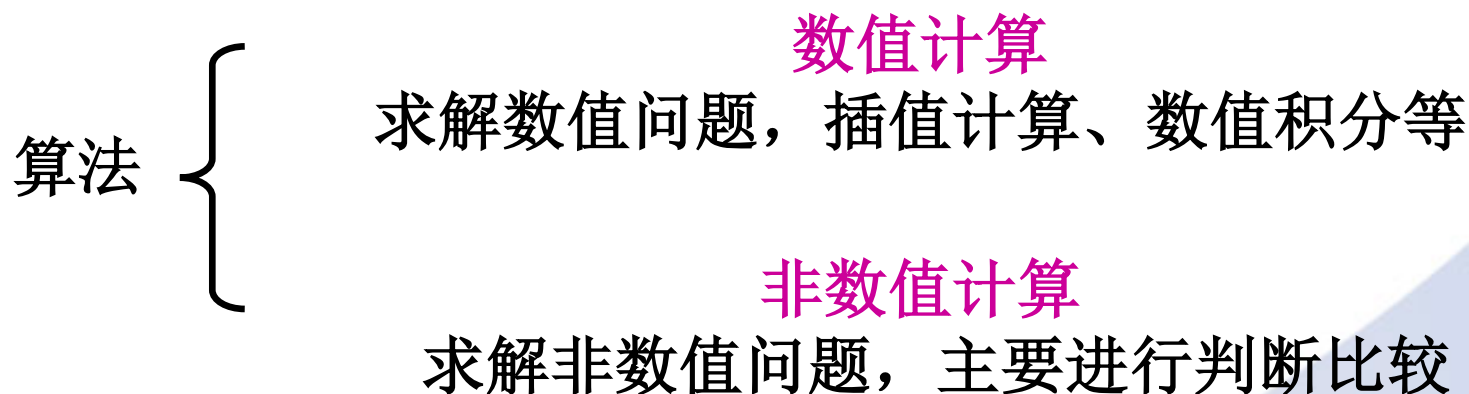
- 1.1 算法的定义及特性
- 1.2 复杂性分析初步
- 1.3 递归



1.1 算法的定义及特性

■ 1. 什么是算法？

- 算法是解一**确定类问题**的任意一种**特殊的方法**。
- 算法是一组**有穷的规则**，它规定了解决某一**特定类型问题**的一系列运算。



- **算法早于计算机出现**



1.1 算法的定义及特性

■ 1. 什么是算法？

□ 算法由运算组成

- 算术运算，逻辑运算，关系运算，赋值运算，过程调用等

□ 算法具有特殊性

- 解决不同问题的算法是不相同的，没有万能的算法

□ 算法是有穷的计算过程

- 静态上，规则/运算/语句的数量有穷
- 动态上，计算过程/计算时间有限



1.1 算法的定义及特性

■ 2. 算法的五个重要特性

□ 确定性、能行性、输入、输出、有穷性

□ 确定性

➤ 算法的每种运算必须要有确切的定义，不能有二义性。

➤ 例：不符合确定性的运算

✓ $5/0$

✓ 将6或7与x相加

✓ 未赋值变量参与运算



中国科学院大学

University of Chinese Academy of Sciences 35

1.1 算法的定义及特性

■ 2. 算法的五个重要特性

□ 确定性、能行性、输入、输出、有穷性

□ 能行性

➤ 算法中有待实现的运算都是基本的运算，原理上每种运算都能由人用纸和笔在有限的时间内完成。

➤ 例：

✓ 整数的算术运算是“能行”的

✓ 实数的算术运算是“不能行”的



1.1 算法的定义及特性

■ 2. 算法的五个重要特性

□ 确定性、能行性、输入、输出、有穷性

□ 输入

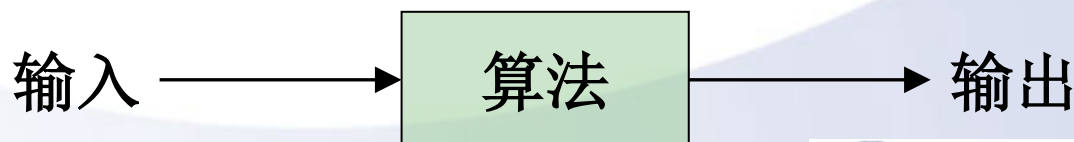
➤ 输入是在算法开始之前给出的量，取自于特定的对象集合——定义域(或值域)

➤ 每个算法有0个或多个输入

□ 输出

➤ 输出是同输入有某种特定关系的量

➤ 一个算法产生一个或多个输出



中国科学院大学

University of Chinese Academy of Sciences 37

1.1 算法的定义及特性

■ 2. 算法的五个重要特性

□ 确定性、能行性、输入、输出、有穷性

□ 有穷性

- 一个算法总是在执行了有穷步的运算之后终止
- 与算法有穷性相关的问题——时效性
- 实际问题往往都有时间要求
 - ✓ 国际象棋(启发)
 - ✓ 数值天气预报
 - ✓ 最优化问题
- 只有在要求的时间内解决问题才是有意义的



1.1 算法的定义及特性

■ 2. 算法的五个重要特性

□ 确定性、能行性、输入、输出、有穷性

□ 程序

- 程序是算法用某种程序设计语言的具体实现。
- 程序可以不满足算法的性质(5)有穷性。
- 任何一种程序设计语言都可以实现任何一个算法
- 有穷性意味着不是所有的计算机程序都是算法
 - ✓ 不能终止的计算过程：操作系统
 - ✓ 算法是“可以终止的计算过程”
- 程序 = 算法 + 数据结构 (By Wirth)



中国科学院大学

University of Chinese Academy of Sciences 39

1.1 算法的定义及特性

■ 3. 算法描述

□ 算法 = 控制结构 + 原操作

□ 表示算法的语言主要有：

- 自然语言
- 流程图
- 盒图
- PAD图
- 伪代码
- 计算机程序设计语言



1.1 算法的定义及特性

■ 3.算法描述

□ 自然语言

- 算法的设计者
- 依靠自然语言交流和表达



人类

机器



中国科学院大学

University of Chinese Academy of Science 41

1.1 算法的定义及特性

■ 3. 算法描述

□ 自然语言

- 自然语言是人们日常所用的语言。使用这些语言，不用专门训练，所描述的算法也通俗易懂。

● 选择排序

- ✓ 第一次遍历找到最小元素
- ✓ 第二次在剩余数组中遍历找到次小元素
- ✓ ...
- ✓ 第 n 次在剩余数组中遍历找到第 n 小元素



1.1 算法的定义及特性

■ 3. 算法描述

□ 自然语言

➤ 自然语言描述算法的缺点：

- ① 自然语言的歧义性易导致算法执行的不确定性。
- ② 自然语言语句一般太长导致描述的算法太长。
- ③ 当算法中循环和分支较多时就很难清晰表示。
- ④ 不便翻译成程序设计语言理解的语言。



1.1 算法的定义及特性

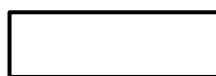
■ 3. 算法描述

□ 流程图

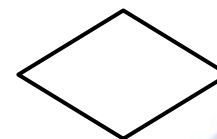
- 流程图是描述算法的常用工具，它采用美国国家标准化协会(ANSI)规定的一组图形符号来表示算法流程图，可以很方便的表示顺序和循环结构。



算法的入口和出口



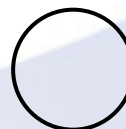
加工、处理



条件



控制流



连接点



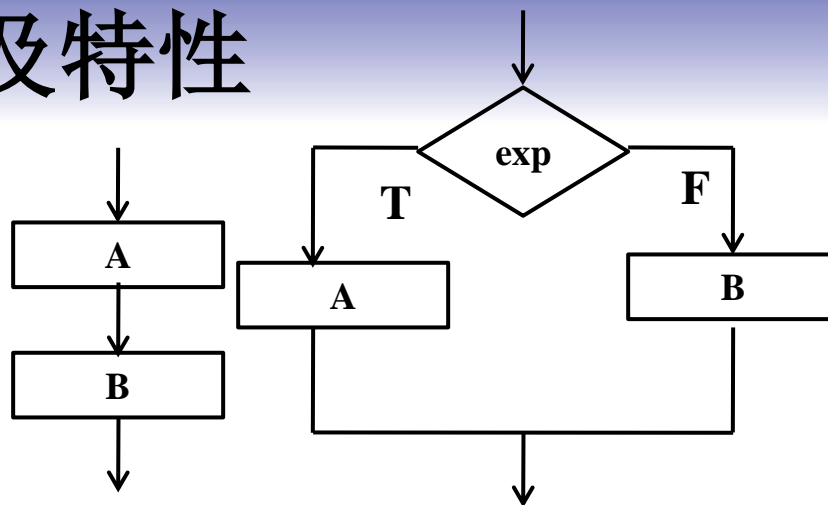
中国科学院大学

University of Chinese Academy of Science 44

1.1 算法的定义及特性

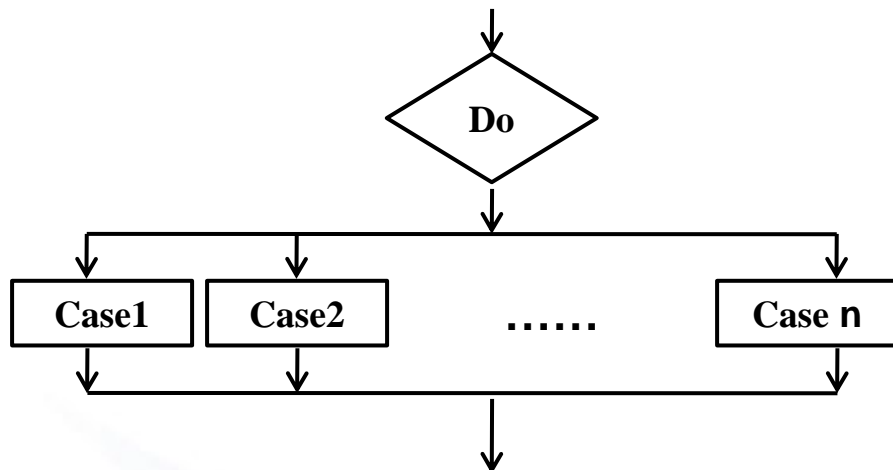
■ 3. 算法描述

□ 流程图

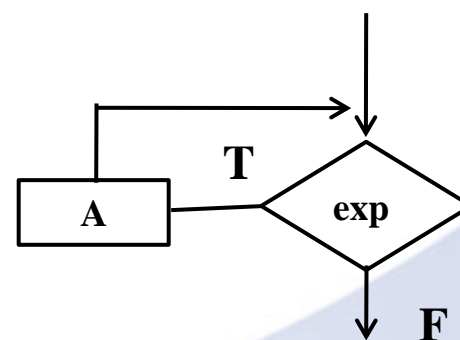


顺序结构

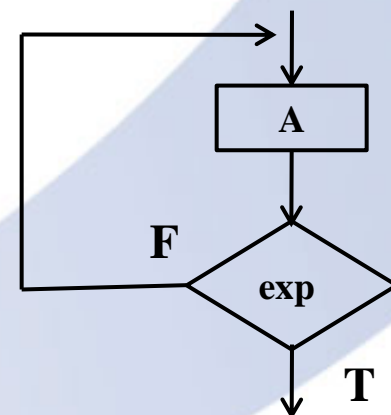
双分支选择结构



多分支选择结构



当型循环结构



直到型循环结构



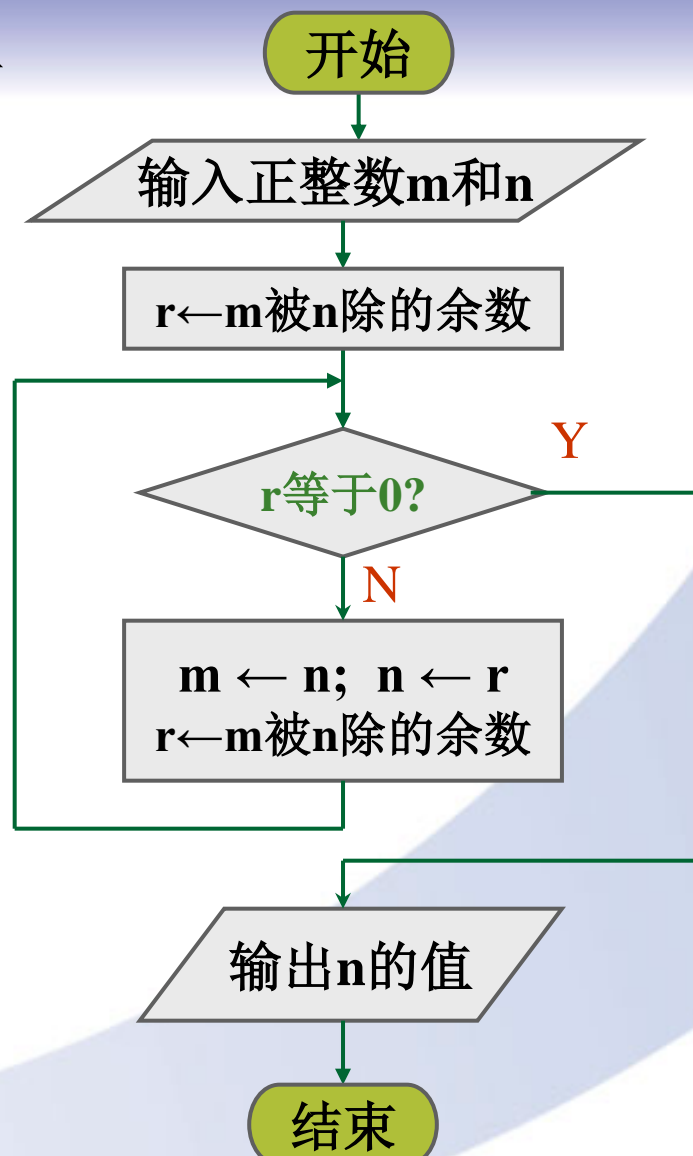
1.1 算法的定义及特性

■ 3. 算法描述

□ 流程图

➤ 主要缺点：

- ① 使算法设计人员过早考虑算法控制流程，而不考虑全局结构，不利于逐步求精。
- ② 随意性太强，结构化不明显。
- ③ 不易表示数据结构。
- ④ 层次感不明显。



1.1 算法的定义及特性

■ 3. 算法描述

□ 程序设计语言

- 算法的执行者
- 需要详细具体的执行代码



人类

机器



中国科学院大学

University of Chinese Academy of Science 47

1.1 算法的定义及特性

■ 3. 算法描述

□ 程序设计语言的缺点：

- ① 算法的基本逻辑流程难于遵循。与自然语言一样，程序设计语言也是基于串行的，当算法的逻辑流程较为复杂时这个问题就变得更加严重。
 - ② 特定程序设计语言编写的算法限制了与他人的交流，不利于问题的解决。
 - ③ 要花费大量的时间去熟悉和掌握某种特定的程序设计语言。
 - ④ 要求描述计算步骤的细节而忽视算法的本质。
 - ⑤ 需要考虑语法细节，而扰乱算法设计的思路。
 - ⑥ 考虑到程序设计语言不断更新，不适于描述算法。
- 算法设计一般不用程序设计语言直接描述。

1.1 算法的定义及特性

■ 3. 算法描述

□ 如何表示一个算法？

- 自然语言：贴近人类思维，易于理解主旨
- 编程语言：精准表达逻辑，规避表述歧义



伪代码



人类

问题：可否同时兼顾两类表示方法的
优势？

机器



中国科学院大学

University of Chinese Academy of Science 49

1.1 算法的定义及特性

■ 3.算法描述

□伪代码

- 伪代码是用介于自然语言和计算机语言之间的文字和符号来描述算法的工具。它不用图形符号，因此书写方便，格式紧凑，易于理解，便于用计算机程序设计语言实现。
- 如：类SPARKS/类C/类C++



1.1 算法的定义及特性

■ 3. 算法描述

□ 伪代码和真实代码的区别在于：

- 在伪代码中，关注的是使用最清晰、最简洁的表示方法来说明给定的算法。可以不用关心**软件工程**的问题。
- 为了更简洁地表达算法的本质，用伪代码描述算法时，常常忽略数据抽象、模块性和错误处理等问题，从而使得算法设计的注意力放在算法**最核心**的部分。

□ 在本课程中，我们将用一种“伪代码”书写程序。伪代码类似于 C, C++, Java, Python, or Pascal，但没有固定的标准



1.1 算法的定义及特性

■ 3. 本课程的主要任务:

□ 算法学习将涉及5个方面的内容:

- ① **设计算法**: 运用一些基本设计策略规划算法
- ② **表示算法**: 用语言把算法描述出来。“类语言”、“伪代码” (Pseudocode, 类SPARKS语言、类C/C++语言)
- ③ **确认算法**: 算法正确性的证明
 - ✓ 算法的正确性: 对合法输入能得出正确的答案。
 - ✓ 算法的证明: 证明算法的正确性, **与语言无关**
 - ✓ 程序的证明: 证明程序的正确性
- ④ **分析算法**: 对算法的时、空特性做定性、定量分析, 以了解算法的性质。



1.1 算法的定义及特性

■ 3. 本课程的主要任务:

□ 算法学习将涉及5个方面的内容:

⑤ **测试程序**: 将算法变成程序, 放到计算机上运行, 观察运行情况

✓ 编程中的**调试**: 排错过程。“调试只能指出有错误, 而不能指出它们不存在错误”

✓ 运行中的**测试**: 分析过程。作时空分布图, 验证分析结论, 进一步优化算法设计。

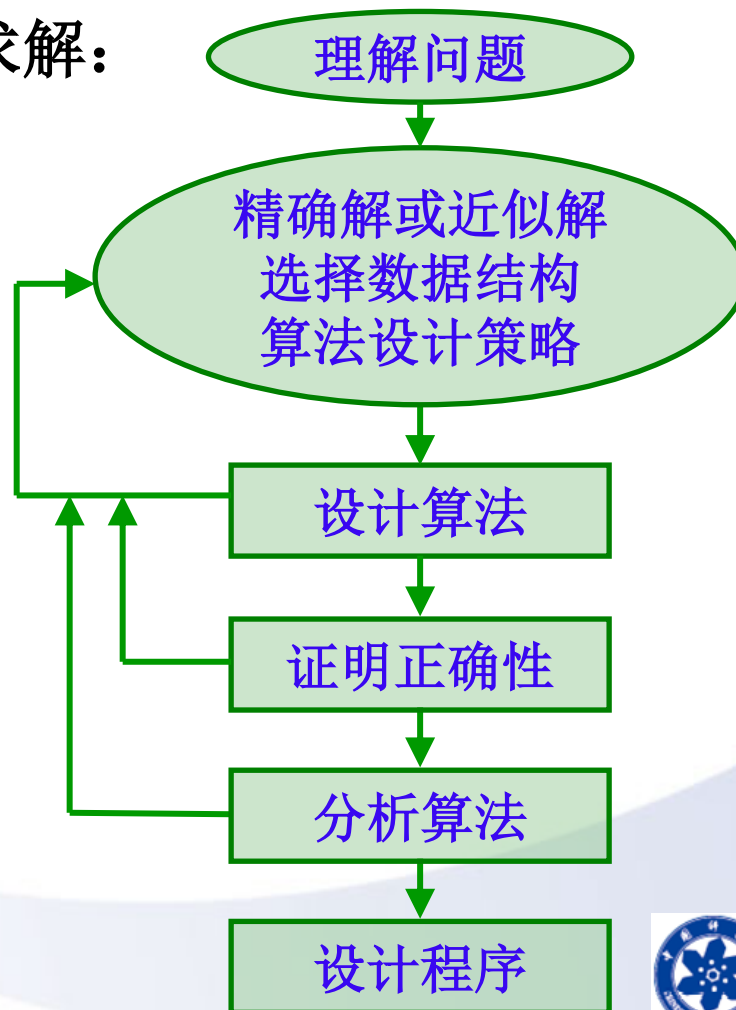
□ 本课程集中于学习算法的**1设计与4分析**。通过学习, 掌握计算机算法设计和分析**基本策略与方法**, 为设计更复杂、更有效的算法奠定基础。



1.1 算法的定义及特性

■ 3. 本课程的主要任务:

□ 问题求解:



End



中国科学院大学
University of Chinese Academy of Sciences