

高等院校信息技术规划教材

# 计算机科学导论

徐志伟 孙晓明 著



清华大学出版社

高等院校信息技术规划教材

# 计算机科学导论

徐志伟 孙晓明 著



清华大学出版社  
北京

## 内 容 简 介

本书从计算思维角度讲解计算机科学最基础的概念和入门知识,讨论计算思维的四种具体表现形式:计算逻辑思维、算法思维、网络思维、计算系统思维。为了足够精准地描述信息变换过程,必须用信息的方式定义并推导信息变换过程涉及的“对”与“错”,哪些能计算,哪些不能计算。这是计算逻辑思维,它往往需要精确地定义计算模型。我们还需要从信息的角度发现和发明解决各类问题的精确方法,并评价什么是有效的方法。这是算法思维。有很多问题不是由单个算法解决,而是由多个算法形成网络来描述和解决。研究有效的网络需要网络思维。信息变换过程往往通过具体的计算设备与系统得以体现。如何设计、评价并使用计算抽象和实用的计算系统涉及计算系统思维。

本书适合作为高等院校计算机及相关专业的教材,也可以作为计算机爱好者的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

计算机科学导论/徐志伟,孙晓明著. —北京:清华大学出版社,2018  
(高等院校信息技术规划教材)  
ISBN 978-7-302-48963-4

I. ①计… II. ①徐… ②孙… III. ①计算机科学—高等学校—教材 IV. ①TP3

中国版本图书馆 CIP 数据核字(2017)第 293490 号

责任编辑:白立军 张爱华

封面设计:常雪影

责任校对:李建庄

责任印制:王静怡

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市铭诚印务有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:18.25

字 数:408千字

版 次:2018年3月第1版

印 次:2018年3月第1次印刷

印 数:1~1500

定 价:45.00元

---

产品编号:064181-01

献给我的母亲：教书育人四十余年的彭老师。

徐志伟

献给我爱人 Tracy 和女儿 Lulu, 感谢你们的支持!

孙晓明

大学计算机基础课程,往往称为“计算机科学导论”,正在经历一场变革,主要体现在三个方面。第一,课程的受众正在扩大。例如,美国著名大学的“计算机科学导论”课程每学期的学生数过去10年增长了50%,一门课程的学生数往往超过700人。第二,课程的内容正在从传统的讲历史发展、讲入门工具使用、讲初阶编程,过渡到讲计算思维。第三,课程的形式正在变得更加丰富,包括课堂讲授、动手动脑实践、慕课、翻转课堂等各种组合。

本书是为中国大学的“计算机科学导论”课程设计的,考虑了全球发展趋势与中国的实际情况,具备下述四个特点。

(1) **强调计算思维**。本书试图突出计算机科学最本质的特征:计算机科学是研究计算过程的科学,计算过程是通过操作数字符号变换信息的过程。最本质的解决问题方法是计算思维,包括逻辑思维、算法思维、网络思维和系统思维。

(2) **强调基础知识**。本书并不追求覆盖众多的时髦名词或新概念,而是突出计算机科学不过时的最基础的知识点,并将它们组织成对计算思维的10个理解:自动执行、正确性、通用性、构造性、复杂度、连通性、协议栈、抽象化、模块化、无缝衔接。

(3) **鼓励主动学习**。本书的设计鼓励同学们自学,但教师讲解与课堂互动有利于揭示要点、提高学习效率。本书还提供了一些动手动脑的大作业和编程练习,对应于逻辑、算法、网络和系统四大内容。

(4) **鼓励扩展眼界**。除了理论内容外,本书提供了较多的实例,涵盖计算机科学及其应用和社会影响。本书花了大约三分之一的篇幅讲解计算机领域的真实的创新故事,让同学们了解前人如何通过计算思维认识世界、提出问题、解决问题。这些创新故事对同学们形成计算机科学领域的学术道德和职业精神也有裨益。

本书的构思与写作持续了五年时间,主要的难点是如何体现计算思维。作者要感谢北京大学李晓明教授,他多年来一直鼓励和敦促我们写一本计算机科学导论教科书。感谢时任美国国家科学基金

会副主任的周以真(Jeannette M. Wing)博士,她多次与我们讨论计算思维的要点。感谢中国科技大学陈国良教授与合肥工业大学李廉教授,以及他们领导的教育部大学计算机基础课程教学指导委员会。这些老师花了很多精力在中国推动计算思维改革,为本书提供了很多经验。特别感谢中国科学院大学的同学们,他们是本书的第一批读者,也是本书作为教科书的“计算机科学导论”课程的第一批实践者。感谢中国科学院计算技术研究所的博士生朝鲁和李春典,他们担任了课程的助教并撰写了课程实践部分的内容。

本书引用了业界的大量素材,在此一并致谢。我们要感谢开源社区,尤其是 LAMP (Linux、Apache、MySQL、Python) 社区。感谢学术社区,尤其是 ACM (Association for Computing Machinery)、IEEE Computer Society、CCF (中国计算机学会)。ACM 与 IEEE Computer Society 是全球最大的计算机科学技术领域的国际学术社区,分别有 10 万与 6 万多名会员。CCF 有 3 万多名会员,是全球第三大计算机学会。

我们还要感谢众多的公司,本书合理使用了它们的素材(例如公司名称、技术和产品名称、logo 标志),这些名称和标志都是这些公司的知识产权。这些公司包括曙光、联想、龙芯、华为、腾讯、百度、IBM、英特尔、谷歌、脸谱网、领英、AT&T、思科、红帽、通用电气、微软、甲骨文、乐高等。免责声明:除了已经公开发表的材料外,本书使用这些公司的例子都做了抽象加工,没有泄露这些公司的隐私。

孙晓明 徐志伟

2017 年夏于北京中关村

<b>第 1 章 计算机科学概貌</b> .....	1
1.1 什么是计算机科学 .....	2
1.1.1 计算思维 .....	4
1.1.2 步骤、符号、操作与计算过程 .....	8
1.1.3 能够自动执行的抽象 .....	16
1.2 计算机科学的发展实例 .....	18
1.2.1 巴贝奇问题——计算机系统实例 .....	19
1.2.2 布什问题——计算机使用模式实例 .....	21
1.2.3 图灵问题——智能应用实例 .....	24
1.2.4 计算机科学的三个奇妙之处 .....	29
1.3 计算机科学的创新故事 .....	36
1.3.1 文王演周易 .....	36
1.3.2 自由软件的故事 .....	41
1.3.3 为什么未来不需要我们 .....	46
1.4 编程练习 .....	49
1.5 习题 .....	53
<b>第 2 章 计算模型与逻辑思维</b> .....	54
2.1 从一个实例看逻辑思维 .....	54
2.2 逻辑思维要点 .....	55
2.2.1 布尔逻辑 .....	55
2.2.2 图灵机模型 .....	62
2.2.3 悖论与不完备定理 .....	66
2.2.4 能够自动执行的精准逻辑 .....	67
2.3 计算逻辑的创新故事 .....	69
2.3.1 布尔的故事 .....	69
2.3.2 图灵的故事 .....	72

2.3.3 臭虫与病毒 .....	73
2.4 习题 .....	79
<b>第3章 算法思维 .....</b>	<b>81</b>
3.1 从一个实例看算法 .....	81
3.2 算法思维的要点与实例 .....	83
3.2.1 分治算法 .....	83
3.2.2 其他算法实例 .....	90
3.2.3 算法复杂度浅介 .....	96
3.3 算法的创新故事 .....	98
3.3.1 平稳复杂度 .....	98
3.3.2 红帽的故事 .....	101
3.3.3 创业公司五步曲 .....	103
3.4 编程练习 .....	111
3.5 习题 .....	113
<b>第4章 网络思维 .....</b>	<b>116</b>
4.1 从一个实例看网络与协议 .....	117
4.2 网络思维的要点 .....	119
4.2.1 名字空间 .....	119
4.2.2 网络拓扑 .....	122
4.2.3 互联网协议栈 .....	124
4.2.4 服务质量与用户体验 .....	126
4.2.5 网络的规律举例 .....	127
4.3 网络的创新故事 .....	130
4.3.1 ARPANET .....	131
4.3.2 以太网、路由器和 TCP/IP 协议 .....	134
4.3.3 万维网 .....	139
4.3.4 社交网络 .....	143
4.4 编程练习 .....	145
4.5 习题 .....	147
<b>第5章 系统思维 .....</b>	<b>150</b>
5.1 从一个实例看计算系统 .....	151
5.2 计算系统思维要点 .....	153
5.2.1 抽象化 .....	154
5.2.2 模块化 .....	159



5.2.3	无缝衔接 .....	172
5.2.4	抽象是应对系统复杂度的利器 .....	179
5.3	计算系统创新故事 .....	180
5.3.1	半导体集成电路 .....	180
5.3.2	计算机硬件 .....	188
5.3.3	计算机软件 .....	213
5.4	编程练习 .....	227
5.5	习题 .....	232
<b>第 6 章</b>	<b>课程实践 .....</b>	<b>234</b>
6.1	构造图灵机实验 .....	234
6.1.1	实验目的和原理 .....	234
6.1.2	实验内容、方法和步骤 .....	235
6.1.3	实验注意事项 .....	236
6.1.4	成绩评定方法 .....	236
6.1.5	思考题 .....	236
6.2	算法游戏实验 .....	236
6.2.1	实验目的和原理 .....	236
6.2.2	实验内容、方法和步骤 .....	237
6.2.3	实验注意事项 .....	238
6.2.4	成绩评定方法 .....	238
6.2.5	思考题 .....	238
6.3	网络路由实验 .....	238
6.3.1	实验目的和原理 .....	238
6.3.2	实验内容、方法和步骤 .....	239
6.3.3	实验注意事项 .....	240
6.3.4	成绩评定方法 .....	241
6.3.5	思考题 .....	241
6.4	淝水之战系统实验 .....	241
6.4.1	实验目的和原理 .....	241
6.4.2	实验内容、方法和步骤 .....	242
6.4.3	实验注意事项 .....	244
6.4.4	成绩评定方法 .....	245
6.4.5	思考题 .....	245
6.5	班级快速排序实验 .....	245
6.5.1	实验目的和原理 .....	245
6.5.2	实验内容、方法和步骤 .....	246
6.5.3	实验注意事项 .....	246

6.5.4	成绩评定方法	247
6.5.5	思考题	247
6.6	莱布尼茨问题实验	247
6.6.1	实验目的和原理	247
6.6.2	实验内容、方法和步骤	247
6.6.3	实验注意事项	248
6.6.4	成绩评定方法	248
6.7	信息隐藏编程实验	248
6.7.1	实验目的和原理	248
6.7.2	实验准备	249
6.7.3	实验内容、方法和步骤	258
6.7.4	实验注意事项	273
6.7.5	成绩评定方法	274
附录 A	计算机科学技术中常用的倍数和分数	275
附录 B	原文阅读列表	277
参考文献		279

## 计算机科学概貌

谁是 20 世纪最伟大的科学家？

由于科学的分支很多，这是一个很难回答的问题。我们可以问更具体的问题。

谁是 20 世纪最伟大的数学家？一个答案是大卫·希尔伯特(David Hilbert)。除了众多的具体数学贡献(如希尔伯特空间)外，希尔伯特还发起和领导了宏伟的希尔伯特计划(Hilbert's program)，试图为数学奠定坚实完备的基础。希尔伯特计划改变了数学学科，对当代数学发展产生了深远影响。

谁是 20 世纪最伟大的物理学家？一个答案是阿尔伯特·爱因斯坦(Albert Einstein)。他提出了光电效应公式、狭义相对论、广义相对论等重要物理理论。最近探测到的引力波，也是爱因斯坦等人最先推测到的。著名的爱因斯坦质能公式  $E=mc^2$  更是广为流传，影响了原子能的研究和利用。

谁是 20 世纪最伟大的农学家？一个答案是袁隆平。他领导研究推出的杂交水稻育种理论和技术，不仅成倍地提高了中国的水稻亩产量，破解了“谁来养活中国”的诘难，还推广到了世界多个国家。

这些伟大的科学工作体现了数学思维、物理思维、生物学中的遗传思维，这些科学思维方式同学们在高中阶段就比较熟悉了。

但是，从本课程角度看，上述问题和答案有一个共性特点：它们并没有体现另外一种日益重要的科学思维方式，即计算思维(computational thinking)。

计算机科学(computer science)是 20 世纪产生的一门相对年轻而又影响广泛的科学。它的最基础、最本质的精髓是一种独特的思维方式，有别于同学们熟悉的数学思维、物理思维、遗传思维方式。随着移动互联网的普及，同学们可能每天都在使用着计算机科学的产品和服务。尽管如此，很多每天使用计算机或手机的同学，对计算机科学的基本思维方式还是会感到陌生。

计算机科学导论这门课程就是介绍计算思维的课程。

## 1.1 什么是计算机科学

计算机科学是研究计算过程的科学。计算过程是信息变换过程(a process of information transformation)<sup>①</sup>。同学们从中学物理就已经知道了物质与能量的运动过程。20世纪中叶以来的一个新发展是:人们已经越来越清晰地认识到,信息变换过程(也就是信息运动过程)与物质运动过程、能量运动过程同等重要。

计算过程是通过操作数字符号变换信息的过程,涉及信息在时间、空间、语义层面的变化。例如一位同学登上长城,将自拍的人物照片用图片文件存档,第二天再打开观看,涉及时间的改变;把人物图片从长城传到海南岛家中给父母看,涉及空间的改变;用一个“老化软件”从人物图片计算出该同学10年后成熟的模样,涉及语义层面的改变。

从技术角度看,计算机科学涉及信息获取、信息存储、信息处理、信息通信、信息显示等环节。一个计算过程可以专注于某个环节,如信息存储过程,也可以覆盖多个环节。为了突出计算机科学这门学科的技术内涵和技术影响,中国业界有时使用“计算机科学技术”(而不是“计算机科学”)来称呼该学科。

### 术语析疑:信息技术

中国业界往往用**信息技术**一词指代“计算机科学技术与通信技术”,用**信息化**一词指代这些技术的应用。国际上则往往采用**信息技术**(information technology, IT)特指“计算机科学技术”,采用一个更大的术语 **information and communication technology**(ICT)指代“计算机科学技术与通信技术”。

在市场上,用户通过三种载体消费信息技术。这三种载体是**硬件产品**(如桌面电脑、智能手机)、**软件产品**(如操作系统软件、数据库软件)、**信息服务**(如信息系统优化服务、互联网视频服务、微信服务)。

从教育角度看,中国的高等教育体系将“计算机科学与技术”设为一级学科,包含三个二级学科:“计算机系统结构”“计算机软件与理论”“计算机应用技术”。国际上的高等教育体系则主要提供“计算机科学”(computer science, 欧洲一些学校使用informatics一词)的本科、硕士、博士学位教育,部分高校提供“计算机工程”(computer engineering)、“信息系统”(information systems)的学位教育。

从人力资源角度看,从事计算机科学技术相关知识、产品、服务的研究、教育、构建和应用的人员往往被称为**信息技术从业人员**,其中具有本科学士及以上文凭的称为**信息技术专业人员**(IT professionals)。他们与信息技术消费者一起所从事的生产活动称为信

<sup>①</sup> “计算机科学”“计算”等本质概念在学术界有多种不断演化的定义。本书的“计算过程”定义主要依据 Richard Karp 教授的定义稍加修改,见 Karp R M. Understanding science through the computational lens[J]. Journal of Computer Science and Technology, 2011, 26(4): 569-577.

息产业(IT industry)。2014年,中国的信息技术专业人员已接近百万人,信息技术从业人员接近千万人,有6亿多信息技术消费者。中国已成为全球信息技术领域用户最多的国家。

计算机专业并不只是培养信息技术从业人员,计算机专业毕业生也并不只是程序员,尽管优秀的程序员一直是稀缺资源。哈佛大学做了一个有趣的调查,看看1984年计算机专业毕业生从事哪些行业。答案(<http://cdn.cs50.net/guide/guide-10.pdf>)是:

- 计算机专业毕业生从事的工作覆盖50多个行业。
- 毕业生从事的工作最多的基本上平均分布在四个行业:教育、工程、信息、科研。
- 毕业生从事的工作第二多的分布在四个行业:金融、工商、政府、传媒。
- 毕业生的工作职务最多的是两个头衔:教授(professor)和总裁(president)。

从应用市场角度看,计算机科学推动了**信息技术市场**的飞速发展。衡量信息市场的一个指标是**信息技术支出**(IT expenditure,或IT spending),含计算机与网络的硬件、软件和服务的每年支出。全球信息技术市场(即全球信息技术支出)1950年仅为数百万美元,1960年增长到数亿美元,2000年达到了1万亿美元,50年期间增长了数十万倍。2000年以后增长放缓,但仍在2013年超过了2万亿美元。

中国信息技术市场起步晚于发达国家,因此增长潜力更大。2000年,中国信息技术市场仅为260亿美元,按全国人口平均计算,人均信息技术支出只有21美元。2008年,中国信息技术市场增长到了1100亿美元,人均信息技术支出提升到89美元。中国科学院在2009年发布了《中国至2050年信息科技发展路线图》研究报告,其中预测:中国信息技术市场将在2040—2050年期间达到每年2万亿美元的规模。这个近20倍增长的一个依据是:如此看似巨大的增长,也只能使中国人均信息技术支出达到1321美元,只是接近美国2000年的水平(1418美元)。根据麦肯锡公司的统计,2013年美国的网民人数已达到2.77亿人,占人口的87%。2013年中国的网民人数仅为6.32亿人,占人口的46%,还有很大的增长空间。

从社会影响角度看,计算机科学已经渗透到人们生产生活的方方面面。学生、劳动者、企业、政府都直接或间接地关注计算机科学,从几岁的小朋友到百岁老人都在使用计算机科学的成果。越来越多的国家已经意识到,人类文明在经历了农业社会、工业社会的发展阶段之后,未来的一个大趋势是进入**信息社会**(information society)。有学者将这些发展阶段称为农业时代(the agricultural age)、工业时代(the industrial age)、信息时代(the information age)。

**实例:**年纪最长的计算机用户。全球年纪最大的计算机用户多少岁?答案是113岁。2014年,美国一位1900年出生的老太太安娜·斯特尔成为脸书(Facebook)用户。她在注册脸书账户时,发现脸书账户注册系统出生时间选项一栏最早只到1905年,被迫谎报年龄才得以注册成功。

**实例:**计算机科学的社会影响。2009年,时任中国国务院总理的温家宝在出席中国科学院建院60周年大会时发表讲话表示,量子力学、相对论、宇宙大爆炸模型、DNA双螺旋结构、板块构造理论、计算机科学,这六大科学理论的突破,共同确立了现代科学体

系的基本结构。

### 1.1.1 计算思维

作为一门专业学科,计算机科学为信息产业提供了主要的知识体系。计算机科学不仅提供一种科技工具、一套知识体系,更重要的是提供了一种从信息变换角度有效地定义问题、分析问题和解决问题的思维方式。这就是作为计算机科学主线的**计算思维**<sup>①</sup>。

可用一句短语概括:**计算思维的要点是精准地描述信息变换过程的操作序列,并使用信息变换过程认识世界、构造性地解决问题。**

事实上,这句短语涉及丰富的内容。中国国家自然科学基金委员会在2010年的一项研究报告中如此刻画计算机科学:计算机科学研究信息的整个生命周期中所有的现象和关系,包括信息的产生、采集、传输、存储、处理、显示和使用。2004年,美国科学院和工程院设立的“计算机科学基本问题委员会”撰写了一部著作<sup>②</sup>,试图总结过去60年的经验,定义计算机科学研究的基本问题。该报告的主要结论如下:“计算机科学是研究计算机以及它们能干什么的一门学科。它研究**抽象计算机的能力与局限,真实计算机的构造与特征,以及用于求解问题的无数计算机应用。**”

报告还总结了计算机科学具有的一些特点:

- 计算机科学涉及符号及其操作。
- 计算机科学关注多种抽象概念的创造和操作。
- 计算机科学创造并研究算法。
- 计算机科学创造各种人工结构,尤其是不受物理定律限制的结构。
- 计算机科学利用并应对指数增长。
- 计算机科学探索计算能力的基本极限。
- 计算机科学关注与人类智能相关的复杂的、分析的、理性的活动。

本书从计算思维这条主线出发,讲解计算机科学的最基础的概念和入门知识,讨论计算思维的四种具体表现形式:**逻辑思维、算法思维、网络思维、系统思维**,如图1.1所示。

组合性	网络思维	系统思维
有效性	算法思维	
正确性	逻辑思维	

图 1.1 计算思维的四种表现形式

<sup>①</sup> 见四个计算思维相关文献:Wing J M. Computational thinking[J]. Communications of the ACM,2006,49(3): 33-35;Xu Z, Li G. Computing for the masses[J]. Communications of the ACM,2011,54(10): 129-137;Karp R M. Understanding science through the computational lens[J]. Journal of Computer Science and Technology,2011,26(4): 569-577;Xu Z W, Tu D D. Three new concepts of future computer science[J]. Journal of Computer Science and Technology,2011,26(4): 616-624.

<sup>②</sup> National Research Council. Computer Science: Reflections on the Field, Reflections from the Field[M]. Washington D. C.: National Academies Press, 2004.

为了足够精准地描述信息变换过程,必须用信息符号的方式定义并推导信息变换过程涉及的“对”与“错”;1 和 0;哪些能计算,哪些不能计算。这是**逻辑思维**,它往往需要精确地定义计算模型(见第 2 章)。**逻辑思维**重点关注计算过程的正确性。需要从信息变换过程的角度,发现和发明求解各类抽象科学问题与应用技术问题的精确方法,并评价什么是有效的方法。这是**算法思维**(见第 3 章)。**算法**重点关注计算过程的有效性。很多问题不是由单个算法解决,而是由多个算法连接组合形成网络来描述和解决。研究有效的网络需要**网络思维**(见第 4 章)。信息变换过程往往通过具体的硬件系统、软件系统、服务系统得以体现,如何设计、评价并使用抽象计算系统和真实计算系统涉及**系统思维**(见第 5 章)。**网络思维与系统思维**重点关注计算过程的组合性,即多个节点如何连接起来组合成为网络,多个模块如何组合成为一个系统。

### 计算思维与编程

计算思维的思想是 2006 年由时任美国国家科学基金会副主任的周以真教授正式提出的。她认为,计算思维就像语文、算术一样,是一种基本的思维能力。到 2050 年,世界上每一个年轻人都应该具备计算思维能力。也就是说,当人类文明从农业社会、工业社会走向信息社会的今天,扫除文盲意味着让每一个孩子都具备语文、算术、计算思维的能力。

计算机科学不仅提供一种科技工具,更重要的是提供了计算思维,即从信息变换角度有效地定义问题、分析问题和解决问题的思维方式。计算机科学不仅是程序设计(俗称编程,programming 或 coding)。那么,程序设计(编程)还重要吗?是否每一个大学生都应该具备基本的编程能力?答案是肯定的。语文并不只是读文章、写文章的能力,但当代青年必须具备基本的读写能力。同理,计算思维并不只是编程,但当代年轻人应该具备基本的读程序、写程序的能力。

英国在这方面走在了世界的前列。2013 年,英国教育部颁布了面向全国中小学生的“计算机程序设计教育国家课程体系”指导意见(见英国政府官方网站 [www.gov.uk/government/publications/national-curriculum-in-england-computing-programmes-of-study](http://www.gov.uk/government/publications/national-curriculum-in-england-computing-programmes-of-study))。意见要求,计算机程序设计课程应该教会学生们(pupils)“使用两种编程语言,其中至少一种是文本语言(即非图形界面语言),解决一系列计算问题”。剑桥大学发明的 Raspberry Pi 廉价计算机(中文翻译为树莓派计算机,售价低至数十美元一台)在这个编程普及教育运动中起到了良好的促进作用。

2016 年 1 月 30 日,时任美国总统的奥巴马公布了称为“全民计算机科学”(computer science for all)的联邦政府计划,拟投资 42 亿美元进一步推动全美的幼儿园至高中(K-12)的计算机科学课程教育。奥巴马认为,计算机科学与读、写、算术一样,是一种基本的能力。

中国政府很早就将计算机教育纳入了中小学教育,这得益于邓小平的名言:计算机教育要从娃娃抓起。今天,“算法”已成为中国中学数学课程的必修单元。

随着计算机科学应用对社会渗透的不断深入,一些新名词变得普遍。其中,“网络空间”“信息空间”“赛博空间”都涉及一个英文词:cyberspace。这是20世纪计算机科学创造的一个新事物。进入21世纪,人们进一步开始认识到,计算机科学事实上涉及人类社会(人)、信息空间(机)和物理世界(物)三者的整体,称为人机物三元世界(human-cyber-physical universe)。这种正在兴起的计算称为三元计算。

下面列出计算过程和计算思维的10种理解。它们各自代表计算机科学的一个本质难题,攻克这些难题成为计算机科学发展史上的里程碑。

### 计算过程和计算思维的10种理解

- 理解1: 自动执行。计算机能够自动执行由离散步骤组成的计算过程。
- 理解2: 正确性。计算机求解问题的正确性往往可以精确地定义并分析。
- 理解3: 通用性。计算机能够求解任意可计算问题。
- 理解4: 构造性。人们能够构造出聪明的方法让计算机有效地解决问题。
- 理解5: 复杂度。这些聪明的方法(称为算法)具备时间/空间复杂度。
- 理解6: 连通性。很多问题涉及用户/数据/算法的连接体,而非单体。
- 理解7: 协议栈。连接体的节点之间通过协议栈通信交互。
- 理解8: 抽象化。少数精心构造的计算抽象可产生万千应用系统。
- 理解9: 模块化。多个模块有规律地组合成为计算系统。
- 理解10: 无缝衔接。计算过程在计算系统中流畅地执行。

如果觉得上述10种理解太多了,还可以从计算机科学的发展历史中进一步提炼出更基本、更简约的目标。从计算机科学技术的发展史看,计算思维的演变大体上出现了五个主要目标体现,并一直延续下来。这五个目标是:自动、通用、算法、联网、抽象。

(1) 自动。这个目标对应理解1:自动执行。计算机科学首先解决的科技难题是将描述物理世界和数学世界的各种公式和方程离散化、数字化,并将连续时间变成离散的步骤,让求解这些公式和方程的计算过程能够自动执行,而不是像使用算盘那样,每一步都需要人工操作。自动执行有助于大幅度提升计算速度,降低计算成本。解决这个难题的历史进程中三个里程碑:1703年莱布尼茨发表二进制算术论文,1837年巴贝奇提出一种称为分析机(analytic engine)的机械计算机的设计,以及1945年ENIAC问世。今天,自动执行难题只能说基本上解决了,并没有得到彻底解决。网络信息空间中业务和技术的异构性、信息孤岛、信息物理系统、人机物三元融合等都为实现自动执行带来了新挑战。

(2) 通用。这个目标对应理解2和理解3:正确性与通用性。在解决了某些数学方程的计算过程自动执行问题之后,人们很自然地猜测:什么样的计算过程能够自动执行?是不是所有的计算过程都能被计算机自动执行?是不是所有求解数学方程的计算过程都能被计算机自动执行?存不存在某一种计算机,它能够执行所有计算过程、解决所有计算问题?通用性难题的一个里程碑发生在1936年,图灵从理论上提出了一种通用计



计算机,能够求解任意可计算问题。而要做到这一点,必须精确地定义并分析计算机求解问题的正确性,包括定义什么是可计算问题、什么是不可计算问题。今天,通用性难题也没有得到彻底解决。例如,即使从理论上,也仍然没有针对如下问题的精确定义与答案:什么是互联网可计算的问题?什么是物联网可计算的问题?什么是众包模式(如人肉搜索)可计算的问题?什么是人机物三元计算系统能够解决的问题?

(3) **算法**。这个目标对应理解4和理解5:构造性与复杂度。知道某个问题可计算还远远不够。一个问题可能通过无穷多种计算过程得到解决。我们需要找到或构造出有效的方法(称为算法),能够花费较短的计算时间、使用较少的计算资源,通过执行比较聪明的特定计算过程来解决问题。要给出算法研究的里程碑比较困难,因为有意义的进展实在太多了,第3章会给出一些实例。随着计算机科学对社会的渗透普及,新的问题层出不穷,针对算法的研究也仍然在蓬勃发展。

(4) **联网**。这个目标对应理解6和理解7:连通性与协议栈。人们很早就注意到了,很多问题只用单点执行的算法不能解决。例如,互联网搜索问题(即搜索数以百万计的互联网网站的信息,找到相关答案),单点算法就不能解决。我们需要将多点的数据和算法连接起来、组成网络,才能解决这些问题。联网难题的进展有很多里程碑,第4章给出了一些例子。

(5) **抽象**。这个目标对应理解8、理解9和理解10:抽象化、模块化和无缝衔接。抽象(abstraction)贯穿计算机科学的发展史,是计算机科学最重要的研究目标和研究对象,同时又是较难掌握的计算机科学的根本方法。计算过程都是在计算系统中执行的。计算机科学并不是针对每一个问题、每一个计算过程设计一套计算系统,而是归纳出少数精心构造的计算抽象,有规律地组合起来,产生万千应用系统。也就是说,计算机科学讲究尽量用一个(或一套)通用抽象支持众多具体应用需求。自动执行、通用、算法、网络都涉及抽象。第5章从系统思维角度详细讨论计算抽象。

### 计算思维是交响乐

《太玄经·差首》说:“帝由群雍,物差其容。”

一种基本的思维方式(帝)往往是多个思想的和谐(群雍),在多样性(物差其容)中涌现出整体之美(帝),就像一首交响乐。它是由多个乐器按照一个乐谱和谐地演奏出来的动听的整体,每个乐器的演奏都发挥出独特的美妙,它们都在表达同一首音乐。

对计算思维,不同的学者有不同的理解、表达和着重点。下面是四个例子。

(1) 加州大学伯克利分校的 Riachard Karp 教授长期倡导算法透镜(algorithmic lens),又称计算透镜(computational lens)概念,强调通过算法思维的透镜观察世界、理解世界,将计算机科学融入自然科学和社会科学。

(2) 牛津大学的 Georg Gottlob 教授认为:“计算机科学是逻辑的继续”(computer science is the continuation of logic by other means),就像“战争是政治的继续”一样。