

计算机科学导论重要概念汇总

数学与应用数学 姚森瀚

*本文的写作得到了陈国凯的大力支持，图片摘自徐志伟老师的课本和幻灯片

0x000F0000

使用 x86 处理器的计算机的 BIOS 的入口地址，它存储 BIOS 的第一条指令。

0xFFFFFFF0

使用 x86 处理器的计算机开机后执行的第一条指令的地址，其内容是一条跳转指令 JUMP 000F0000。

ABC 特征 (ABC features)

计算思维的外部特征，是自动执行、比特精准、抽象构造首字母的简写。

ACM 行为准则 (ACM code of conduct)

计算机协会官方网站称为《计算机协会道德与职业行为准则》。在本课程中，特指其第一部分：一般道德原则，计算机协会提出的七条原则：

- (1) 为社会和人类的幸福做出贡献，承认所有人都是计算的利益相关者；
- (2) 避免伤害；
- (3) 诚实可靠；
- (4) 做事公平，采取行动无歧视；
- (5) 尊重需要产生新想法、新发明、创造性作品和计算工件的工作；
- (6) 尊重隐私；

(7) 尊重保密协议。¹

根据 ACM 行为准则，发布 Morris 蠕虫的行为是错误的，因为他违反了准则的第二条：避免伤害。

ALOHANET

亦称 **ALOHA 系统**、**ALOHA 网络**。由美国夏威夷大学所开发的实验性电脑网络系统。在 1971 年 6 月开始成功运作，是世界上第一个无线资料网络²。ALOHA 网络使用 **ALOHA 协议** 工作。

ALU (arithmetic and logic unit)

亦称**运算器**、**算术逻辑单元**、**算术逻辑部件**。**CPU** 的组成部分。能实现多组算术运算和逻辑运算的组合逻辑电路。

ASCII (American Standard Code for Information Interchange)

亦即**美国信息交换标准代码**。一套编码系统，用一个字节给字符编码。因为最左端的一位总是 0，所以 ASCII 字符的值在 0~127 之间。其中：

0~31 对应的是 NUL~US；32 对应的是 SP；33~47 对应的是 ! ~ / (一些符号)；

48~57 对应的是 0~9；58~63 对应的是 :~? (一些符号)；64 对应的是 @；

65~90 对应的是 A~Z；97~122 对应的是 a~z；127 对应的是 DEL。

¹ 复制自计算机协会的官方中文网站。

² 论述来自百度百科。

阿姆达尔定律 (Amdahl' s law)

简单来说，阿姆达尔定律断言，“系统性能的改善受限于其瓶颈部分。”

假如一个系统可以分成两部分 X 和 Y， $X + Y = 1, 0 \leq X, Y \leq 1$ 。Y 能够被改善，X 不能被改善，则系统最多能够被改善到 $1/X$ 。

这里的“被改善到”是按照加速比的意义来理解的。

暗文 (ciphertext)

密码学术语。加密后的文本。

巴贝奇问题 (Babbage' s problem)

如何造一台计算机？更确切地说，如何建造有效的、可编程的计算机？

巴贝扬断言/巴贝扬黄金譬喻 (Babayan' s gold metaphor)

俄罗斯科学院院士波瑞斯·巴贝扬认为，计算速度是黄金一样的硬通货，可以换成其他任何东西，包括新的功能、易用性、更好的产品、成本更低的服务，等等。

保底异常 (machine check)

亦称机器检查，一种异常。通过硬件方式，覆盖其他规定的异常没有覆盖的情况，做到穷尽所有可能。

包体 (packet body)

包的有效载荷数据³。见“包交换”。

包交换 (packet switching)

亦称**分组交换**。贯彻整个**互联网协议栈**的通信原则，即通过“**包头+包体**”的消息包传递信息。每个层次传输的一条消息被分解成很多小的“**包**”，都包含两个部分：一部分是用户真正想传输的**载荷信息**（又称**包体**，**负载**），另一部分是用作控制或管理的其他信息，称为该消息的**首部**（又称**包头**，**元数据**）。可以类比写信时信封和信纸的关系来理解这个过程。分组交换将多个用户的多条消息拆分成数据包，依次随机发送。

包头 (packet header)

包含控制信息、地址、错误处理信息等。见“包交换”。

贝尔定律 (Bell' s law)

计算机的发展遵循三种设计方式，每十年左右产生一种新的计算机类型。

- (1) 发展性能最好的计算机，将价格作为次要的考虑因素；
- (2) 在控制价格的基础上，改善该计算机的性能；
- (3) 尽可能地将价格降低，生产最低价格的计算机。

迄今为止市场上出现了大约 10 种计算机类型，分为两大类：

- (1) 多用户服务端计算机；
- (2) 单用户客户端计算机。

³ 参见英文教材 p265.

比特 (bit)

即二进制位。

边

网络拓扑概念。指节点间的连接。

编译时错误 (compile-time error)

编译器在编译程序时发现的程序错误。常见的错误有括号不匹配、没有定义、没有赋初始值、类型不匹配、程序与环境不匹配。

病毒 (virus)

即计算机病毒。恶意软件的一种。一种可以自我复制并且可以“感染”计算机的程序。计算机病毒需要嵌入另一个程序中（“感染”该程序）。

病毒性市场 (viral marketing)

在网络计算时代，在计算机的全民普及阶段，产品和服务需要具有病毒的特征才能最快速、最广泛地流行。为什么计算机病毒易于流行呢？专家们总结了病毒的六个市场特征：连通性、幂数律、低价格、好使用、易传播、强黏性。

标志位 (flag bit)

见 FLAGS。

并行计算 (parallel computing)

使用多个处理器执行同一个计算任务的方法。

并行计算机 (parallel computer)

并行计算的计算机。

波纹进位加法器 (ripple-carry adder)

将数个全加器串行级联起来得到的加法器，进位变量的值像池塘中波纹一样逐级传递。它真实地反映了人们用纸和笔做加法的过程。

博斯特尔健壮性原理 (Postel' s Robustness Principle)

亦称宽进严出原理。理解无缝衔接需要的四条原理之一。该原理的一个目的是避免误差、漂移的积累。

布尔常数

两个具有恒定值的符号：0 和 1。可以看作特殊的布尔变量⁴。

布尔函数 (Boolean function)

$$f: \{0,1\}^n \rightarrow \{0,1\}$$

两个布尔函数相同，当且仅当它们有相同的真值表。

⁴ 论断来自中文教材。

n 个变量的布尔函数有 2^{2^n} 个。

布尔函数的范式定理

任何一个布尔表达式可以被等价地变换成下面的唯一表达式：它是若干积的和，而在每一个积中， x_i 或 \bar{x}_i 刚好出现一次（常量 0 是 0 个积的和）。

这个定理的直接推论是，任何布尔函数都可以用三级门来实现。可以想象，至少在理论上，今天的计算机的速度还可以大大提高。

布尔逻辑运算

即二值逻辑运算。它是最基本的运算操作。

补码 (two's complement)

表示计算机中有符号数的一种方法。正数的补码是其二进制表示，与原码相同。

负数将其原码除符号位的所有位取反后加一。

有符号二进制数的加减法可通过补码实现，它的原理是模算术。

例如，8bit 表示下-10 的补码可以这样计算：

先写出-10 的有符号表示：10001010（原码）；

然后讲其除符号位以外的所有位取反：11110101（反码）；

最后加 1 得到 11110110（补码）。

布什问题 (Bush' s problem)

如何使用计算机？更确切地，如何将计算机更方便地、更有效地运用在解决问题

的过程中？布什问题关注计算机的使用模式。

布唐蜜蜂譬喻 (Boutang' s bee metaphor)

信息通信技术就像蜜蜂，生产两种价值。直接价值是蜂蜜，间接价值（经济外部性）是授粉。布唐估计，授粉的经济价值要远高于蜂蜜。相类似地，数字经济的价值远高于信息通信市场本身的价值。

CMOS (Complementary Metal Oxide Semiconductor)

即**互补金属氧化物半导体**。实现半导体电路的主流技术。它是指制造大规模集成电路芯片用的一种技术或用这种技术制造出来的芯片。

CPU (central processing unit, CPU)

即**中央处理器**，简称**处理器**。

操作

基本的操作可分为三类：**运算操作、存储操作、输入/输出操作**。

层间接口

在一个设备中协议栈相邻两个层次间的接口。

超链接 (hyperlink)

也称**超级链接**。某超文本文档中的元素与因特网上另一个元素之间的链接。

超文本 (hypertext)

也称**超级文本**。超文本是用**超链接**的方法，将各种不同空间的文字信息组织在一起的网状文本。

程序计数器 (program counter, PC)

一个特殊的**寄存器**，用于存放将要执行的下一条指令的地址。详见“指令集抽象”。
在执行一般指令时， $PC=PC+8$ （假设指令是 64 位字长，即 8 个字节）。
在执行跳转指令时，PC 被赋予跳转目的地址的值。

程序周期 (program cycle)

程序的一次执行时间从时间上的开始到结束构成一个程序周期，它是通过多条指令的执行构成的，执行的机制是**指令流水线机制**。

臭虫 (bug)

即**计算机臭虫**。用来指称计算机或计算机网络的所有错误，是指计算机系统的硬件、系统软件（如操作系统）或应用文件（如文字处理软件）出错，包括**软件臭虫**和**硬件臭虫**。臭虫属于“无心之失”，即非故意产生的错误。

抽象 (abstraction)

抽象化的产物。

抽象化 (abstraction)

系统思维的关键概念⁵之一（同时也是计算思维的内部特征之一），是系统思维最本质的考虑。即从多个层次（角度，视野）理解一个系统，每个层次仅考虑有限的、该层次特有的问题，忽略其他问题；并用一套精确规定的抽象概念和方法，统一地处理该层次所有的计算过程，解决这些特有问题。

抽象化的产物称为**抽象**。抽象化和抽象满足**抽象三性质**。

计算机学科的抽象化特色是可自动执行的、比特精准的信息抽象，即**数据以及对数据的操作**。**数据抽象**和**控制抽象**是计算机科学的抽象特色的重要体现。

抽象三性质(Three Properties of Abstraction; COG properties of abstraction)

抽象化和抽象满足的三个性质：**有限性、精确性、通用性**。

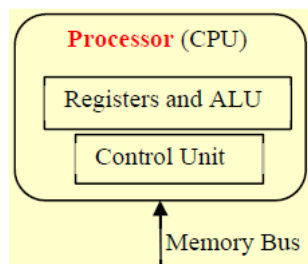
触发器 (flip-flop)

又称门闩，由逻辑电路和反馈线组成。

处理器 (processor)

亦称 CPU，冯·诺依曼体系结构中**硬件**的组成部分。它在一个**控制单元**的控制下，使用 **ALU** 和少量**多用途寄存器**执行指令。**ALU、寄存器和控制器**一起组成的基本的中央处理器。

⁵ 英文版课本称这三者是“主要特征” (main character)，同时也称这三者是“关键概念” (key concept)，我们采后者。



现代 CPU 可能还包括其他的处理单元，例如图形处理、机器学习处理。

传输层 (transport layer)

TCP/IP 协议栈的第四层，一种广泛使用的协议是 **TCP**，它在两台互联网主机间发送 **TCP 包**。

传输控制协议 (Transmission Control Protocol, TCP)

即 **TCP**。

串行执行 (serial execution)

冯·诺依曼体系结构的要点之一。一个计算过程是一个串行执行过程，任何一个程序都是通过一条接一条指令的**自动执行**来执行的。

存储操作

对数据的操作的一种。典型的存储操作包括将数据存储在某地方，或者从某地方取出来放在寄存器中。这个“地方”的名字称为存放该数据的**地址**。

存储电路

亦称“状态电路”。

存储器^① (memory)

亦称内存储器、主存储器，简称内存、主存。冯·诺依曼体系结构的硬件组成部分之一，处理器可以通过指令访问它。寄存器可以被看成处理器中特殊的存储单元。

存储器^② (storage)

亦称存储设备。它是一种 I/O 设备。典型的例子有硬盘驱动器 (HDD)、固态硬盘 (SSD)。

DDoS 攻击 (DDos attack, distributed denial-of-service attack)

利用多台互联网主机累积攻击总量的 Dos 攻击。

Dos 攻击 (Dos attack, denial-of-service attack)

亦称拒绝服务攻击。指短时间内发送大量的信息，导致计算机无法调用足够的资源去处理正常的请求。常见的 DoS 攻击有带宽攻击和连通性攻击。⁶
更进一步地，有 DDoS 攻击。

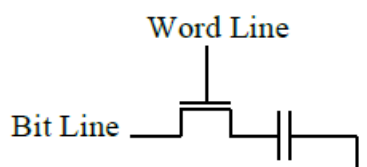
DNS (Domain Name System, DNS)

⁶ 这一句补充自互联网资料。

又称**域名系统**。它将**域名**和它的**IP 地址**联系起来。

DRAM (dynamic random access memory)

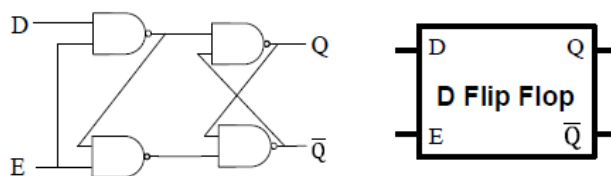
亦称**动态随机存取存储器**，一种**易失性存储器**。



它由一个电容和一个晶体管构成，这种简单的构造使得它成本低廉。然而电容是会漏电的，这也使得 DRAM 需要每 $7.8-128 \mu s$ 就刷新自己的内容。

D 触发器 (D flip-flop, delay flip-flop)

亦作**延迟触发器**。D 触发器由四个与非门和三个反馈回路构成，它的行为可以通过真值表来刻画。当使能信号 E 是 0 时，D 触发器保持它的状态，即 $Q_{next} = Q$ ；当使能信号 E 是 1 时，D 触发器把它的状态改变成 D 的输入值，即 $Q_{next} = D$ 。



实际常用**时钟信号**来充当使能信号，此时 D 触发器输出的状态相比输入值 D 要延迟一个时钟周期，这也是它得名的原因。

带宽 (bandwidth)

信息通过网络传递的**比特速率**。

单调布尔函数 (monotonous Boolean function)

记 $(x_1, \dots, x_n) \leq (y_1, \dots, y_n)$, 如果 $x_i \leq y_i, i = 1, 2, \dots, n$.

若 $(x_1, \dots, x_n) \leq (y_1, \dots, y_n) \Rightarrow f(x_1, \dots, x_n) \leq f(y_1, \dots, y_n)$, 称布尔函数 f 是单调的, f 是单调布尔函数。

0 元单调布尔函数有 2 个, 1 元单调布尔函数有 3 个, 2 元单调布尔函数有 6 个, 3 元单调布尔函数有 20 个。

滴漏效应 (trickle down effect)

在科学研究高端领域产生的研究成果会滴漏到更加量大面广的企业计算乃至消费者计算领域。

这是因为科学研究领域对计算的需求比较大, 对计算的投入也比较高, 而且强调自由探索和受控试验, 很多创新成果会首先在科学研究领域出现并得到验证。

第一原理研究法 (study from first principle)

纯粹从公理出发研究一个问题的方法。

地址 (address)

① 存放数据的位置的名称。见“存储操作”。

② 网络术语。地址是一种特殊的名字, 可以被用于直接访问一个实体。非地址名字应当先被转换成一个地址, 然后才能访问该名字指称的实体。

电信服务 (telecommunication service)

指利用有线、无线的电磁系统或者光电系统等各种通信网络资源，提供语音通话服务，传送、发射、接收或者应用图像、短信等电子数据和信息的业务活动。例如，电话业务、互联网连接业务。

钓鱼 (phishing)

亦称网络诱骗。通过一定的“诱饵”引导用户给出自己的敏感信息。

钓鱼有钓鱼网站和钓鱼邮件两种方式。

动态网络 (dynamic network)

按网络拓扑变化分类的一种网络。动态网络不改变自己的节点，但可能改变自己的边。

对称加密 (symmetric-key encryption)

密码学术语。在这种加密过程中，发送方与接收方共享密钥。

一个经典的例子是凯撒密码。

对等接口⁷ (Peering Interface)

同一层次的通信双方交互的接口。对等接口屏蔽了底层的细节，增强了易用性。

对数据的操作 (operation on data)

⁷ 英译来自计科导 PPT。

包括存储操作、运算操作、通信操作。

多媒体计算模式 (multimedia mode)

按人机交互方式分类的一种使用模式。计算机不仅支持图形，还支持图像、声音、视频等。

多用途寄存器 (general-purpose register)

见“寄存器”。

恶意软件 (malware)

特意设计的可以帮助攻击者破坏或未获得授权地进入目标计算机的软件。常见的包括病毒、蠕虫、木马、间谍软件。

二进制位 (binary digit, bit)

亦称比特。二进制记数系统的位，后也用于表示两种状态。

二进制指数退避 (binary exponential backoff)

一种指数退避。当第一次传输的试图失败之后，计算机等候从 $0 \sim T$ 的一个随机时间值；第二次重试失败后，等候 $0 \sim 2T$ 的一个随机时间值；第三次重试失败后，等候 $0 \sim 4T$ 的一个随机时间值；……以此类推。理论和实验证明，采用了这个方法，任何计算机都有很大的概率不会无穷地等待。

FLAGS

一种程序状态寄存器。它持有一组标志位，用于标记一条指令的执行状态，例如返回值的符号、是否有溢出等。

泛化 (generalization)

抽象的一种性质，见“通用性”。

反向滴漏效应 (trickle up effect)

进入 21 世纪，计算机科学技术领域出现了一个新的现象：创新成果首先在量大面广的个人消费者计算领域出现，然后扩散到企业计算与科学研究领域。

防病毒软件 (antivirus software)

即杀毒软件。

访存操作

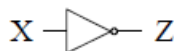
CPU 访问存储器的取数存数操作。

防火墙 (firewall)

防范网络犯罪和网络攻击的一种方法。防火墙安装于互联网与计算机之间，阻挡或筛除不想要的信息。

非门 (NOT gate)

一种门。



非易失性存储器 (non-volatile memory, NVM)

断电后仍能保存其内容的存储器。有两种常见的类型：只读存储器 (ROM) 和可读写非易失存储器 (read-write NVM)。

分治算法 (divide and conquer)

将求解一个大的问题归约成求解一个或多个规模较小的子问题递归求解。

例如，单因素优选法。

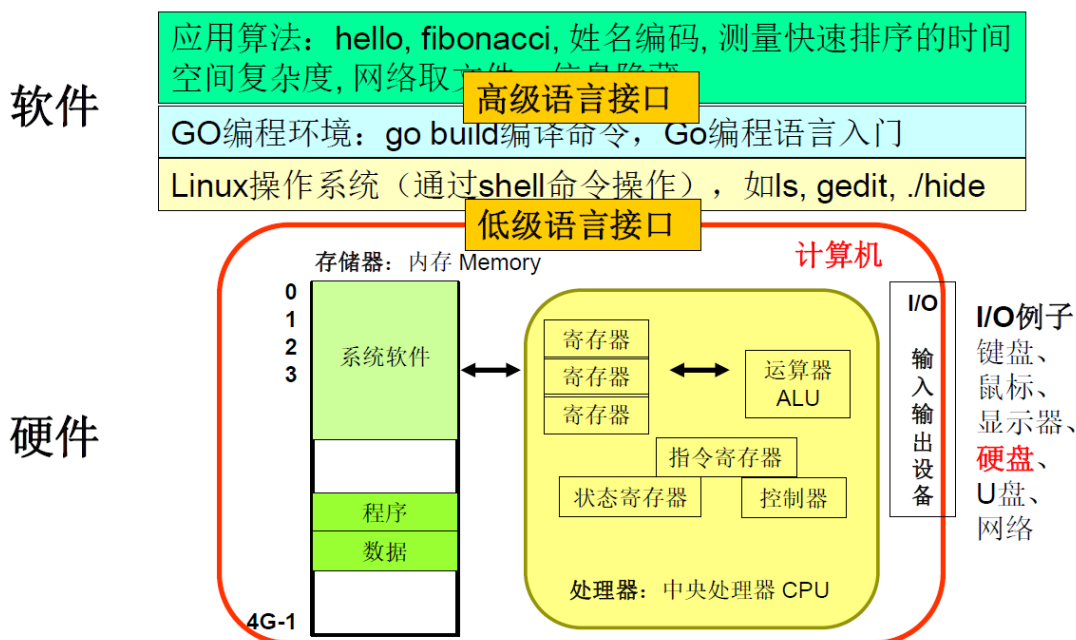
冯·诺依曼穷举原理 (von Neumann' s Exhaustiveness Principle)

理解无缝衔接需要的原理之一。要使计算机自动执行程序，人们必须事先给计算机全面的指示，绝对穷举所有细节，使得计算机能够自动处理所有的情况，执行过程中不需要人工干预。这些细节包括计算机需要执行的程序的指令、程序的输入数据、程序需要的函数等。后来人们发现，还需要规定好如下实现细节：

- (1) 计算机开机后的第一条指令；
- (2) 计算机正常执行程序时，下一条指令是什么；
- (3) 执行程序出现异常时，有哪些异常，每种异常如何自动处理。

冯·诺依曼体系结构 (von Neumann architecture)

又称冯·诺依曼结构、冯·诺依曼模型，存储程序计算机体系结构。



当代计算机基本组成: **冯诺依曼模型简介**

12

其要点是:

- (1) 二进制数据及其算术逻辑操作。
- (2) 计算机包含处理器 (运算器、控制器、寄存器)、存储器、输入/输出设备。注意处理器和存储器是分开的两个单元。它们之间的连接 (内存总线) 后来被称为计算机的冯·诺依曼瓶颈。
- (3) 存储程序计算机。这里又分为普林斯顿体系结构、哈佛体系结构两种。
- (4) 指令驱动的串行执行。

浮点数 (floating point number)

一种计算机表示实数的方法。最广泛使用的标准是 IEEE 754 浮点数标准。浮点数形如 $\pm M \times 2^E$ 。浮点数不可以用 == 来比较, 而应当用两者差的绝对值与一个小

的阈值去比较。

符号 (symbol)

指代具体或抽象事物的特定记号，其表现形式可以是文字、数值、图像、声音等。

服务端计算机 (server-side computer)

在机房里边的计算机。用户通过**客户端计算机**间接使用服务端计算机，如各种服务器、超级计算机。

服务器 (server)

接受并响应**客户端**请求的主机。

Gale-Shapley 算法 (Gale-Shapley algorithm)

针对**稳定匹配问题**提出的一种算法。

算法分成若干轮进行。开始的时候算法把每一位女生都标记为“自由的”。在第一轮中每一位男生从自己的偏好序列中选取排在第一位的女生，并向该女生提出共舞的邀请。如果一名女生原来是“自由的”，但是在这一轮中收到了至少一名男生的邀请，则将她标记从“自由的”改成“不自由的”。

如果没有任何一名女生是“自由的”，由于男生与女生的数量相同，因此此时每一名女生恰好收到一名男生的邀请，所有这些邀请构成一个完美匹配。那么她们接受这一邀请，算法输出这一匹配并结束。

如果还有“自由的”女生，算法执行下面的步骤。

由于有一些女生是“自由的”，即没有任何邀请，那么一定有女生收到了多于一名男生的邀请。对于每一名收到多于一名男生邀请的女生，她将选择所有邀请他的男生中在她自己的偏好排序中最高的一位作为她当前计划共舞的对象，同时她将拒绝掉所有其他邀请。如果她的标记是“自由”的，则将此标记改成“不自由的”。对所有那些遭到女生拒绝的男生，他们将从自己的偏好队列中选取尚未被自己邀请过的、排位最高的女生，并向其发出邀请，不管其当前是否有计划共舞的对象。算法重新检查是否存在“自由的”女生，如果有，则重复上述步骤。

高速缓冲存储器⁸ (cache)

简称**高速缓存**，介于**处理器**和**内存**之间的一种规模较小、速度很快的存储器。它自己被组织为两个层次：第一层包括分立开的**指令高速缓冲存储器**和**数据高速缓冲存储器**，每一个都持有 32KB 的信息，但和 CPU 一样快；第二层是一个单独的高速缓冲存储器，它被第一层的两个存储器共享，存储 256KB 信息，但是比第一层更慢。

高速缓存 (cache)

即**高速缓冲存储器**。

个人计算模式 (personal computing mode)

按人机交互方式分类的一种**使用模式**。让每个用户独占一台计算机，不受他人干扰。典型的例子是个人计算机。

⁸ 汉译来自互联网高频用法。

格式占位符⁹

格式占位符(%)是在 C/C++ 语言中 (golang 大体相似) 格式输入函数, 如 scanf、printf 等函数中使用。其意义就是起到格式占位的意思, 表示在该位置有输入或者输出。格式占位符的形式如下:

% [标志][输出最小宽度] [.精度] [长度] 类型

按类型分类, 常见的格式占位符有:

%c 读入一个字符 (得名于 character 的首字母)

%d 读入十进制整数 (得名于 decimal 的首字母)

%i 读入十进制, 八进制, 十六进制整数 (得名于 integer 的首字母)

%o 读入八进制整数 (得名于 octal 的首字母)

%x, %X 读入十六进制整数 (得名于 hexadecimal 中间的 x)

%s 读入一个字符串, 遇空格、制表符或换行符结束。(得名于 string 的首字母)

%f, %F, %e, %E, %g, %G 用来输入实数, 可以用小数形式或指数形式输入。

%p 读入一个指针 (得名于 pointer 的首字母)

%u 读入一个无符号十进制整数 (得名于 unsigned 的首字母)

%n 至此已读入值的等价字符数 (得名于 number 的首字母)

构造性¹⁰ (effectiveness)

直译为有效性, 计算思维的**内部特征**之一。指人们能够构造出聪明的方法让计算

⁹ 内容补充自 CSDN。

¹⁰ 汉译来自计科导 PPT。

机有效地解决问题。

广播 (broadcast)

网络术语。一个设备通过**总线**向其他所有设备发送信息。

归约

设 A 和 B 是两个可计算问题，称可以从问题 A 归约到问题 B，如果任给一个求解问题 B 的算法，都可以“使用”此算法求解问题 A，记作 $A \leq_p B$ 。

HTTP (hypertext transfer protocol, HTTP)

即**超文本传输协议**。一种应用层协议，通常运行于 TCP 之上，指定了客户端可能发送给服务器什么样的消息以及得到什么样的响应。

HTTPS (Hyper Text Transfer Protocol over Secure Socket Layer, HTTPS)

即**超文本传输安全协议**。在 HTTP 的基础上通过传输加密和身份认证保证了传输过程的安全性，加入了 SSL。

哈佛体系结构 (Harvard architecture)

计算机采用两个存储器：一个是数据存储器，一个是指令存储器。高速缓冲存储器采用了哈佛体系结构。

汉诺塔 (Hanoi tower)

亦称**河内塔**，源于印度古老传说。大梵天创造世界的时候做了三根金刚石柱子，在一根柱子上从下往上按照大小顺序摞着 n 片黄金圆盘。大梵天命令婆罗门把圆盘从下面开始按大小顺序重新摆放在另一根柱子上。并且规定，在小圆盘上不能放大圆盘，在三根柱子之间一次只能移动一个圆盘。僧侣们预言，当所有的金片都从梵天穿好的那根针上移到另外一根针上时，世界就将在一声霹雳中消灭，而梵塔、庙宇和众生也都将同归于尽。

经过简单的推理可以证明，三柱汉诺塔最少的移动步数是 $2^n - 1$ 。

核 (core)

并行计算机中的每个**处理器**称为核。

合取范式 (conjunctive normal form, CNF)

布尔函数的一种范式。

$$f(x_1, \dots, x_n) = Q_1 \wedge \dots \wedge Q_m$$

其中， $Q_i = l_1 \vee \dots \vee l_n, l_j = x_j$ 或 $\neg x_j$ 。

有定理指出，有 n 个变元的命题 $F(x_1, \dots, x_n)$ 一定可以表示成合取范式。

互联网计算模式 (network computing mode)

按人机交互方式分类的一种**使用模式**。用户可以通过**客户端计算机**上网。开始主要是通过桌面电脑、笔记本电脑和工作站上网。后来衍生出**云计算模式**。

互联网主机 (Internet host)

连接到互联网的计算机，这里的计算机包含连接到互联网的设备。

环境¹¹ (environment)

包括人（用户）和输入/输出设备。

幻数 (magic number)

编程者在写源代码的时候，使用了一个数字，他当时是明白这个数字的意思的，但是别的程序员看他的代码，可能很难理解，甚至，过了一段时间，代码的作者自己再看代码的时候也忘记了这个数字代表的含义。虽然不知道这个数字是干什么用的，究竟代表什么，但是编译后的程序可以正常运行，真是“魔术般的数字”，“幻数”因此得名。幻数的可读性差，而且不便于修改，因此在编程中应尽量避免使用幻数。

回环 IP 地址 (loopback IP address)

127.0.0.1 这个 IP 地址为本地计算机保留下来，称为回环 IP 地址。

回环域名 (loopback domain name)

localhost 这个域名为本地计算机保留下来，称为回环域名。

回应机理

¹¹ 英译是编者自己判断的。

BBN 公司为解决 ARPANET 的可靠性问题提出的方案。当节点 A 的 IMP 向节点 B 传送一条消息时，它会一直存储着完整的信息，直到节点 B 的 IMP 收到这条消息并送回一个确认信息。如果在一段预先规定好的时间里节点 A 的 IMP 没有收到确认信息，它会认为传输失败，再重新传送同样的信息。

或门 (OR gate)

一种门。



霍克尼公式 (Hockney' s formula)

描述延迟和带宽关系的公式。

$$t = t_0 + \frac{m}{r_\infty}$$

其中 t 指传送信息的总时间， t_0 是最小延迟， r_∞ 是最大带宽。

用户体验到的延迟是 t ，体验到的带宽是 $\frac{m}{t}$ 。特别地，空消息的传输时间也不是 0。

该公式可以解释很多现象，例如用户购买了 1-Gbps 带宽，但只享受到 5-8Mbps 的服务，可能是因为用户传输了大量的短消息，导致实际体验到的带宽远低于最大带宽。当然，也有可能是因为与其他用户共享这 1-Gbps 带宽，或只有接入端是 1-Gbps 带宽。

IP 地址 (IP address, Internet Protocol address)

互联网的每个节点需要被命名，至少应当有一个地址，这个地址称为 IP 地址。分

为 IPv4 地址和 IPv6 地址两种。IPv4 有 32 位地址格式，IPv6 有 128 位地址格式。

基础软件 (infrastructure software)

软件栈的组成部分，由中间件（上层）和系统软件（下层）组成。

寄存器 (register)

本课程中仅考虑了两大类寄存器，即多用途寄存器和专门用途寄存器，后者又分为状态寄存器 FLAGS 和程序计数器 PC。

机器智能 (machine intelligence, AI)

亦称人工智能。研究图灵问题的领域。

机群 (cluster)

多个计算机互联而成的一套计算机系统。

计算过程 (computational process)

计算过程是通过操作数字符号变换信息的过程，是信息变换过程，涉及信息在时间、空间、语义层面的变化。

计算机病毒 (computer virus)

即病毒。

计算机臭虫 (computer bug)

即臭虫。

计算机科学 (computer science)

研究计算过程的科学，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。中国业界有时使用“计算机科学技术”来称呼该学科。

中国国家自然科学基金委员会在 2010 年的一份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，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。”

计算机模拟 (computer simulation)

简称模拟，又称计算机实验。利用执行计算机程序的方法来模仿物理或社会过程。

计算机蠕虫 (computer worm)

即蠕虫。

计算机实验 (computer experiment)

即计算机模拟。

计算思维 (computational thinking)

计算思维是一种从信息变换角度认识世界并改造世界的思维方式，其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。

计算思维的四种具体表现形式是**逻辑思维**、**算法思维**、**网络思维**和**系统思维**。其中，逻辑思维重点关注计算过程的**正确性**，算法重点关注计算过程的**有效性**，网络思维与系统思维重点关注计算过程的**组合性**。

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**。

计算思维的外部特征强调**比特精准**、**抽象构造**、**自动执行**，又称**ABC 特征**。

计算思维的内部特征是**自动执行**、**正确性**、**通用性**、**构造性**、**复杂度**、**抽象化**、**模块化**、**无缝衔接**、**连接性**和**协议栈**，又称**Acu-Exams-CP**。

“计算思维是交响乐。”一种基本的思维方式往往是多个思想和和谐，在多样性中涌现出整体之美，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。¹²

计算系统思维

¹² 论述来自中文教材。

简称**系统思维**。计算系统思维的要点是：通过**抽象**，将**模块组合成为系统**，**无缝执行计算过程**。构造计算系统是有条理、有门道的 (systematic)，不是随意而为 (ad hoc)，但系统思维同时要求创造性和综合性。

这套系统思维方法经过 70 年的发展卓有成效，但也仍有许多不足之处。

研究、理解和使用计算系统的最大挑战是应对系统的复杂性。破解系统复杂性挑战的主要思维方法是抽象化。

计算系统思维包括**抽象化**、**模块化**、**无缝衔接**三个关键概念。

不管是怎样的系统，都需要思考以下三个本质的问题：

- (1) 系统需要什么样的抽象？
- (2) 系统由哪些模块如何组合而成？
- (3) 系统的部件如何无缝衔接，流畅地执行计算过程？

集线器 (hub)

一种**组网设备**。一个小的**局域网**往往是通过将一些**主机**连接到一个称为**集线器**的中心设备实现的**总线结构**。

机械执行 (mechanical execution)

即**自动执行**。

架构师 (architect)

计算机系统设计者。

加密 (encryption)

密码学术语。将明文转变为暗文的过程。

加速比 (speedup)

用于描述改善前后系统执行时间的缩短，定义为改善前后系统的执行时间之比。

简单命题 (simple proposition)

即原子命题。

间谍软件 (spyware)

恶意软件的一种。这种软件收集被入侵计算机的信息，并报告给攻击者。

交换机 (switch)

一种组网设备。交换机连接多台主机或多个局域网，通过这种方式可以组成一个更大的局域网。

交互式计算模式 (interactive computing mode)

按人机交互方式分类的一种使用模式。用户与计算机即刻交互。例如，用计算机生成一个 3000 字符的文本文件，用户一边用键盘输入字符，一边就马上看到屏幕上的字符显示。

节点 (node)

连接而构成网络的部件。

网络中的一个节点可以是一台计算机，也可以是一个硬件部件、一个软件服务、一个数据文档、一个人或一个物理世界中的物体。

接口消息处理器 (interface message processor, IMP)

第一个网络交换机。ARPANET 的每个节点计算机不是与其他节点计算机直接相连，而是通过一台称为接口消息处理器的小计算机连上 50kb/s 的电信专线而连通。计算机 A 想与计算机 B 通信，则先把消息传给它的 IMP，转换成一种 IMP 之间能够理解的格式，传给计算机 B 的 IMP，后又翻译成计算机 B 能够理解的格式传给计算机 B。

解密 (decryption)

密码学术语。将暗文恢复为明文的过程。

接入点 (access point, AP)

一般指无线接入点，又称无线访问节点。一种组网设备，它将无线信号转换为有线信号，反之亦然。AP 设备是 WiFi 局域网的一部分。

进程 (process)

一台计算机上通常会有多种应用程序运行。当代计算机采用一种机制，称为进程机制，来统一执行这些应用程序。每个程序在运行时都被看作一个进程，由操作

系统按照同样的方式统一管理、执行。每个进程都有同样的结构（程序区、数据区、栈区、堆区），都有同样的生命周期阶段（诞生、就绪、运行、睡眠、死亡等）。进程是在执行中的程序，是动态的，而程序本身（软件代码）是静态的。

进程机制

见“进程”。

精简指令集计算机¹³（Reduced Instruction Set Computing）

一种执行较少类型计算机指令的微处理器，它能够以更快的速度执行操作。

精确性

抽象三性质之一。抽象化的产物是一个计算抽象，它是一个语义精确、格式规范的计算概念。

静态网络（static network）

依网络拓扑变化方式分类的一种网络，指节点和边不改变的网络。

典型的例子是星型网络。

局域网（local area network, LAN）

把本地的多台计算机互联起来的网络。

¹³ 英译来自百度百科。

KISS 原理 (KISS Principle)

计算系统设计的一条原理。在设计一个计算系统时，一定要保持简单、傻瓜 (Keep It Simple, Stupid)。

卡普计算透镜论题 (Karp' s computational lens thesis)

很多自然和人类社会的过程也是计算过程，自然界和人类社会都在进行某种计算。把它们看做计算过程，可以给我们带来新的视角、新的价值。这种方法也叫“透过计算透镜” (through the computational lens)。

开放源码 (open source)

泛指种种保证了自由软件的许可证，这些许可证的区别是它们规定了不同的限制，如是否可以包含专有软件部件、修改的部分是否可以变成专有、是否可以重新定义软件许可证，例如 GPL，LGPL，FreeBSD，公有软件等。

可读写非易失性存储器(read-write non-volatile memory, read-write NVM)

一种非易失性存储器。可以读出，也可以再次写入。

一个常见的例子是学生 USB 记忆棒中的闪存。

客户端 (client)

直接被用户使用的主机。与之相对的概念是服务器。

客户端计算机 (client-side computer)

人们最熟悉的计算机，它们直接被人使用。包括各种桌面电脑、笔记本电脑、平板电脑、智能手机等。

科克定律 (Keck' s law)

唐纳德·科克观察了近 40 年的光纤传输的破纪录实验，指出单根光纤的数据传输速率随时间指数增长，大约每 10 年增长 100 倍。

科学计算

科学计算本质上是面向科学家的计算，主要用户是科学家和工程师，主要计算负载是求解方程，包括代数方程、常微分方程、偏微分方程。

科学文献图 (scientific literature graph, SLG)

所有的科学文献构成一个图，称为科学文献图。论文或著作构成该图的节点，引用构成了从引用作品指向被引作品的边。

控制抽象 (control abstraction)

通常针对数据抽象的多个操作步骤组合起来才能解决一个问题。控制多个步骤如何组合起来实现计算过程的操作抽象称为控制抽象。

三种常见的控制抽象是：顺序、条件跳转、调用。

库米定律 (Koomy' s law)

斯坦福大学的乔纳森·库米博士研究了自从 1945 年 ENIAC 诞生至 2010 年的计算机性能功耗比数据，发现了一个趋势：计算机性能功耗比随时间指数增长，大约 1.57 年翻一番。

但是，根据库米等人最新的研究，计算机性能功耗比仍然随时间指数增长，但现在大约 2.7 年才翻一番。中科院计算所研究了高性能计算机近 70 年来的发展历史，观察到类似的危机：计算机系统性能功耗比的提升大幅放慢，滞后于性能的提升速度。

快速排序算法 (quicksort)

计算机常用的排序算法，其核心思想是不断递归调用自身来对子问题进行排序。伪代码如下：

QuickSort (A,p,r)

If $p < r$

1. $q = \text{Partition}(A,p,r)$

 //从数组 $A[p,\dots,r]$ 中随机抽出一个元素 x ，然后对数组进行调整，把比 x 大的数都排在 x 的右边，比 x 小的数都排在 x 的左边，暂不用进行左右两边大小关系的排序，最后返回 x 在数组中的位置 q 。

2. $\text{QuickSort}(A,p,q-1)$

3. $\text{QuickSort}(A,q+1,r)$

快速排序算法的期望运行时间是 $O(n \log n)$ ，它显著地快于冒泡排序算法。

垃圾邮件 (spam)

用户不想要的邮件称为**垃圾邮件**。

里德定律 (Reed' s law)

里德定律的表述很简单：

$$\text{网络价值} = 2^{C-1}$$

其中 C 是网络中的“群组”或“社区”个数。

Google 的工程师将里德定律拓展为

$$\text{互联网价值} = 2^W$$

其中 W 指代万维网。

目前，里德定律还缺少实证支持。

离散无穷性 (discrete infinity)

见“乔姆斯基数字无穷性假说”。

连接性 (connectivity)

计算思维的**内部特征**之一。很多问题涉及用户/数据/算法的连接体，而非单体。

路由器 (router)

一种**组网设备**。当我们想要连接多个使用不同协议的网络时，就要用到路由器。

路由器可以连接两个**异构网络**。

逻辑连接词 (logic connective)

简单命题可以通过连接词组成复合命题。常见的连接词包括与、或、非、蕴含、异或。

MAR (memory address register, MAR)

亦称**内存地址寄存器**，持有用于访问内存的地址。

MDR (memory data register, MDR)

亦称**内存数据寄存器**，持有用于保存写入内存或者从内存读出的数据。

梅特卡夫定律 (Metcalfe' s law)

网络的价值与网络节点数的平方成正比。另一种说法是，网络的价值与上网人数的平方成正比。

2013-2015 年期间，梅特卡夫本人和中科院计算所的最新研究考察了脸书公司与腾讯公司的十余年的历史数据，都嫩被梅特卡夫定律所刻画，这说明梅特卡夫定律有一定的普适性。

门 (gate)

从逻辑设计角度看，计算机系统最基本的抽象是**布尔逻辑门电路**，简称门。

五种常见的门是与门、或门、非门、异或门、与非门。

门闩 (latch)

即**触发器**。

密码学 (cryptography)

研究在攻击者存在的条件下进行安全信息通信的学科。¹⁴

密钥 (key)

密码学术语。一段特别设计的信息，用于加密、解密和完整性验证。

命名系统

即名字空间。

命题 (proposition)

非真即假的陈述句，其真假称为真值。命题分简单命题和复合命题。

命题变量 (propositional variable)

亦称命题变元。数学上为了符号化，用字母表示命题，称为命题变量。

命题逻辑 (proposition logic)

推理由命题变量和逻辑连接词组成的复合命题的逻辑系统。

明文 (plaintext)

¹⁴ 定义来自英文版课本。

密码学术语。加密前的原始文本。

名字 (name)

名字是一个用于标志实体的字符串。

名字冲突 (name conflict)

名字在网络空间中不唯一的现象。

名字空间 (name space)

亦称命名系统，主要用于规定网络节点的名字及其合法使用规则，也可包括命名其他客体（如消息、操作）等的规则。

设计与理解名字空间时，最基本的考虑是名字空间能够指称网络的所有节点，另外还有一些基本考虑：

- (1) **名字的唯一性**：名字是否在全网唯一？
- (2) **名字的自主性**：用户能自主地确定和修改某个网络节点的名字吗，还是需要某种权威机构确定？
- (3) **名字的友好性**：名字是否容易理解和使用？
- (4) **名字解析**：相关但不同的两个名字空间如何对应并自动地翻译？

因特网有两个最重要的名字空间：一个是**域名空间**，一个是**IP 地址空间**。网络中最最重要的一个名字空间注册工作就是互联网域名注册绑定到一个 IP 地址。

摩尔定律 (Moore' s law)

半导体芯片中晶体管的数量随时间指数增长，约每两年翻一倍。

人们发现半导体芯片技术的研究开发有一种自我加强性，它本身也不断地用到它产生的新型半导体芯片产品，从而进一步改进自身。人们还提出了一种数学公式反映这种加强性：

$$\frac{dTech}{dt} = kTech$$

模块化 (modularity/modularization¹⁵)

系统思维的关键概念之一，即理解每一个系统是由多个模板按一定规则连接组合而成的。也可以反过来理解：一个系统如何分解成多个模块的组合。

模块化可用来应对复杂性。理解一个系统的过程可以通过理解每一个模块以及这些模块的组合规则来简化。事实上这一点大有讲究，是计算机领域的重要创新环节。人们将“系统由多个模块组合而成”的成功的方法称为**系统架构模型**，并赋予特定的名称。

模块化方法需要回答三个问题，它们合起来称为**系统架构三问题**。

有时模块也被称为**子系统**。

模拟之妙¹⁶ (wonder of simulation)

计算机科学发展历史上的三大奇妙之一。**计算机模拟**成为理论和实验之外的第三种科学探索方式。

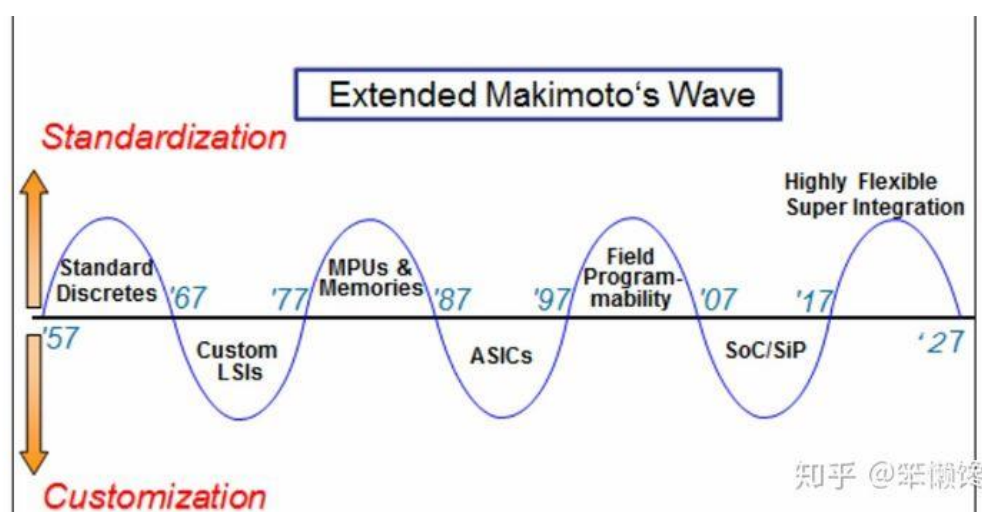
¹⁵ modularity 来自中文版课本，modularization 来自英文版课本。

¹⁶ 汉译来自中文课本。

牧本浪潮 (Makimoto's wave)

牧本浪潮是一个与摩尔定律类似的电子行业发展规律，它认为芯片有规律的在“专用”与“标准、通用”之间变化，循环周期大约为 10 年。这些周期同时受芯片制造技术和芯片的市场需求两类因素推动产生。

牧本浪潮中的新周期的出现并不意味着现有浪潮的消失，而是意味着新市场的涌现。牧本次生预测 2017-2027 年是另一个通用化阶段，其特征是 HFSI。



木马 (Trojan horse)

亦称特洛伊木马，恶意软件的一种。正如希腊故事流传的那样，这种程序将它们的意图隐藏起来，伪装成常见的计算机文档。它表面上是一段合法的程序或数据，但隐藏在它之中有一段通常比较短小的恶意程序。

NP 问题 (NP problem)

可以在非确定型图灵机上在多项式时间内解决的问题。另一种说法是，在确定型图灵机上可以在多项式时间内验证解是否正确，但不能在多项式时间内找出最优

解的问题。

逆否命题 (contrapositive proposition)

亦称质位变换命题。命题 $P \rightarrow Q$ 的逆否命题是 $\neg Q \rightarrow \neg P$ ，它和原命题是等价命题。

诺德豪斯定律 (Nordhaus' s law)

计算速度随时间指数增长。从 1945 年至 2006 年，计算速度每年增长 50%。

O 记号 (big O notation)

假设 f 和 g 都是从非负整数映射到非负整数的函数，如果 f 和 g 满足：存在一个常数 $c > 0$ ，使得对于任意的 n ， $f(n) \leq cg(n)$ ，那么称 $f(n) = O(g(n))$ 。

o 记号 (small o notation)

假设 f 和 g 都是从非负整数映射到非负整数的函数，如果 f 和 g 满足：存在一个常数 $c > 0$ ，使得对于任意的 n ， $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$ ，那么称 $f(n) = o(g(n))$ 。

Ω 记号 (Omega notation)

假设 f 和 g 都是从非负整数映射到非负整数的函数，如果 f 和 g 满足：存在一个常数 $c > 0$ ，使得对于任意的 n ， $f(n) \geq cg(n)$ ，那么称 $f(n) = \Omega(g(n))$ 。

P 问题 (P problem)

确定型图灵机可以在多项式时间内解决的问题。

批处理模式 (batch processing mode)

以人机交互方式为分类依据的一种使用模式。用户将计算任务（包括计算程序与输入数据）提交给计算机。计算机花几秒、几小时、几天甚至几个月的时间执行计算任务。完成计算任务后，计算机将输出结果返回给用户。

普惠计算

面向个人消费者需求的计算机科学研究与应用。

普林斯顿体系结构 (Princeton architecture)

计算机包含一个统一的存储器，可存储数据与程序。

企业计算

面向企业的计算，主要用户是企业的员工、管理者、上下游合作伙伴和客户，主要负载是 workflow、事务处理、数据分析、决策支持等。有时政府部门也被看成企业，电子政务是企业电子业务的一个特例。还有一类特殊的企业计算称为嵌入式计算。

嵌入式计算

一类特殊的企业计算。包括企业生产现场的计算机实时控制、工控系统、机器人等。

嵌入式计算机 (embedded computer)

这样的计算机装在**嵌入式系统**里边，看起来不是计算机。

嵌入式系统

指内部含有计算机控制的自动或半自动系统。

乔姆斯基数字无限性假说¹⁷ (Chomsky' s digital infinity principle)

亦称“离散无穷性”。一组有限的数字符号可以组合创造无限的有意义的表达。换句话说，任何领域内的问题和知识都可以被该领域的专业语言描述，这种专业语言又可以被数字符号描述，从而可以用计算机处理。

丘奇-图灵论题 (Church-Turing hypothesis)

任何人类用纸和笔所能做的计算与图灵机能做的计算等价。也就是说，图灵机刻画了人类所有“可计算”的问题。

在计算复杂性理论出现之后，学者们又提出了更强的丘奇-图灵论题：任何人类能做的计算都可以被图灵机有效地模拟。这里有效的意思是指“多项式时间”归约。丘奇-图灵论题是否正确目前仍不清楚。

取指 (instruction fetch, IF)

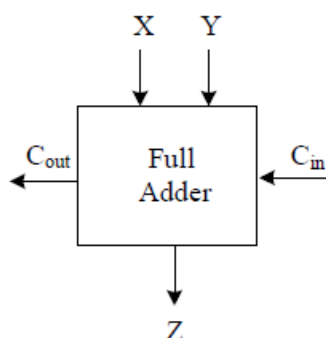
指令流水线四个操作阶段之一。取指就是“取出指令”的缩写。取指前，这条指

¹⁷ 这个译名来自计科导 PPT.

令放在**程序计数器 (PC)** 里。取指的过程是，通过指令的**地址**在**内存**中读取指令，将当前要执行的指令读取到处理器中，放在**指令寄存器 (IR)** 里。

全加器 (full adder)

考虑进位的一位加法器。它可以用下面的**全加器符号**表示：



其中， X 、 Y 是这一位的输入变量， C_{in} 是从上一位来的进位输入， Z 是这一位的结果输出， C_{out} 是这一位的进位输出。

全加器的布尔表达式如下：

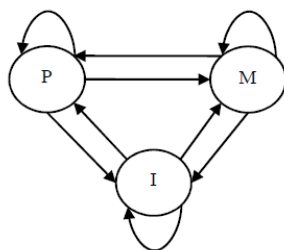
$$S = X \oplus Y \oplus C_{in}$$

$$C_{out} = (X \cdot Y) + ((X \oplus Y) \cdot C_{in})$$

全双工¹⁸ (full-duplex)

网络术语。通信允许数据在两个方向上同时传输。

¹⁸ 汉译来自互联网广泛使用的说法。



ROM¹⁹ (read-only memory)

亦称只读存储器、固定存储器，一种非易失性存储器（NVM）。只能读出无法写入信息。信息一旦写入后就固定下来，即使切断电源，信息也不会丢失。ROM 所存数据通常是装入整机前写入的，整机工作过程中只能读出，不像随机存储器能快速方便地改写存储内容。

人工智能 (artificial intelligence, AI)

亦称机器智能。研究图灵问题的领域。

人机交互方式 (style of human-computer interaction)

使用模式关注的问题之一。用户与计算机的主要交互方式是什么？

以人机交互方式为例，历史上层层出现了下列使用模式，都还被使用着，包括批处理模式、交互式计算模式、个人计算模式、图形用户界面模式、多媒体计算模式、便携式计算模式、互联网计算模式、移动互联网计算模式。

人机物三元世界 (human-cyber-physical universe)

¹⁹ 内容摘编自网络。

人类社会、信息空间和物理世界三者的整体。进入 21 世纪后，计算机科学事实上已经涉及到该三领域的整体。

容错性 (fault tolerance)

针对分布式系统，容错性是指在系统出现一定错误的情况下，系统仍可正确有效运行的性质。

熔毁²⁰ (meltdown)

一种利用硬件的攻击，发现于 2018 年。这种攻击“不利用任何软件的弱点”，而是利用了硬件的乱序执行，使得攻击者可以读取有权限的信息。

蠕虫 (worm)

即计算机蠕虫。恶意软件的一种。一种可以独立自我复制并传播到其他电脑的软件。

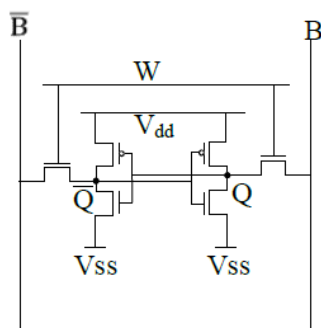
软件栈 (software stack)

通常计算机上的软件都以一种层级结构组织起来，称为软件栈，它在冯·诺依曼体系结构的上层。软件栈由应用软件和基础软件构成，基础软件处于底层、更基本的位置。

SRAM (static random access memory)

²⁰ 汉译来自互联网广泛使用的说法。

亦称静态随机存取存储器，一种易失性存储器。它由六个晶体管组成，比 DRAM 要贵得多，但是也要快很多。



S-R 门闩 (S-R latch)

置位-复位门闩的简称。一个仅含两个与非门的触发器电路。

该电路有四种可能的行为：

- 保持状态： $S = R = 0, Q_{next} = Q$
- 置位： $S = 1, R = 0, Q_{next} = 1$
- 复位： $S = 0, R = 1, Q_{next} = 0$
- 禁止： 要避免出现 $S = R = 1$ 的情况

三副本容错

通过三副本与三选二表决器实现容错计算。假设需要计算 $y = F(x)$ ，而系统会出错。一个经典的思路是三副本容错，即用三个系统分别计算 F ，得出三个结果 A, B, C ，再通过一个“三选二表决器”布尔逻辑电路，计算出最后的结果。

三元计算

涉及人机物三元世界的计算。

杀毒软件 (antivirus software)

又译**防病毒软件**。用于侦测和消灭计算机病毒的软件。事实上，现在的杀毒软件也可以侦测和消灭其他**恶意软件**。

舍入误差 (rounding error, roundoff error)

进行数值计算的过程中，对计算得到的中间结果数据要使用“四舍五入”或其他规则取近似值，因而使计算过程有误差。它不是一种溢出错误²¹。

例如，计算机计算 $2.0/7.0$ 时，便有可能得到舍入误差。

时间共享 (time sharing)

亦称**分时**。指多个用户分享使用同一台计算机。多个程序分时共享硬件和软件资源。时间共享技术是**交互式计算**的关键技术。

时序电路 (sequential circuit)

由**组合电路**与**状态电路**组合产生的电路。

本课程中的时序电路实际上是**时钟同步时序电路**，因为状态电路是由时钟信号驱动的²²。

²¹ 论断来自计科导大作业。

²² 这一论断来自大作业 5。

使用模式²³ (usage mode)

使用模式关注用户群、信息组织方式、人机交互方式三个问题，可以简写为

使用模式=用户群+信息组织方式+人机交互方式

使用模式和**网络思维**密切相关，用户群往往意味着一个用户网络，信息组织方式往往是信息互联方式，人机交互方式是人与计算机之间的互连方式。

计算机的使用模式直接影响计算机市场，同时，每一种使用模式都需要相应的技术支撑，包括硬件、软件、算法等方面的技术进步。

时钟频率 (clock frequency)

一秒内**时钟周期**的数量。

时钟信号 (clock signal, CLK)

在 0 (LOW) 和 1 (HIGH) 之间变换的信号，其循环周期称为**时钟周期**。

时钟周期 (clock cycle)

时钟信号的概念。0 (LOW) 和 1 (HIGH) 的一个循环周期称为**时钟周期**。

手工计算 (manual computing)

人用笔和纸做计算，以及借助于算盘等工具做计算。

数据 (data)

²³ 英译来自英文版教材。

指一个或多个**数字**符号。

数据抽象 (data abstraction)

亦称**数据类型**、**数据结构**。指将某一类数据及其操作的合称，见“对数据的操作”。比较常见的七种数据抽象：**比特**、**4 比特**、**字节**、**字**、**指针**、**文件**、**数据**。它们也体现了基本上从小到大的七个抽象层次。

数据放置问题

如何放置数据，使得读取数据的速度最快、硬盘空间最省？本课程介绍了中科院计算所何永强的新技术 RCFfile，现已被广泛使用。

数据结构 (data structure)

见“数据抽象”。

数据类型 (data type)

见“数据抽象”。

数据链路层 (data link layer)

TCP/IP 协议栈的第二层，它在两种同构连接的设备间传递**帧**。两种广泛应用的协议是**以太网**（由 IEEE 802.3 标准明确）和**WiFi**（由 IEEE 802.11 标准明确）。

数据压缩 (data compression)

减小文件大小的技术，用以节省存储空间、减少传输时间。

根据是否可从压缩文件恢复原文件，可以分为**有损压缩**和**无损压缩**。

输入/输出设备 (Input/Output device, I/O device)

亦称 I/O 设备，冯·诺依曼体系结构中硬件的组成部分之一，包括**硬盘**、**键盘**、**鼠标**、**显示器**、**打印机**等。

数字电路

组合电路和**时序电路**合起来称为**数字电路**。

数字化

将数值**离散化**也称**数字化**²⁴。

数字 (digit)

某种离散的计数方法中的基本数值，也是最简单的一类**符号**。

例如，在 $6+9=15$ 这个加法运算中，6、9、1、5 都是数字。

数字符号 (digital symbol)

能用一个或多个**数字**的组合表示的**符号**。

²⁴ 解释来自计科导中文教材。

例如，在 $6+9=15$ 这个简单的加法运算中，6、9、1、5 是数字符号，15、+、= 也是数字符号。

算法 (algorithm) ²⁵

一个算法是一组有穷的规则，给出求解特定类型问题的运算序列，并具备下列五个特征：

- (1) **有穷性**：一个算法在有限步骤后必然要终止；
- (2) **确定性**：一个算法的每个步骤都必须精确地（严格地和无歧义地）定义；
- (3) **输入**：一个算法有零个或多个输入；
- (4) **输出**：一个算法有一个或多个输出；
- (5) **能行性**：一个算法的所有运算必须是充分基本的，原则上人们用笔和纸可以在有限时间内精确地完成它们。

算法分析

算法研究中的“发现”工作，指找出算法的内在规律。

算术逻辑部件 (arithmetic and logic unit, ALU)

即 ALU。

Θ 记号 (Theta notation)

假设 f 和 g 都是从非负整数映射到非负整数的函数，如果 f 和 g 满足： $f(n) =$

²⁵ 这是引用高德纳教授的算法定义。

$o(g(n))$ ，且 $f(n) = \Omega(g(n))$ ，则称 $f(n) = \Theta(g(n))$ 。

TCP (Transmission Control Protocol, TCP)

即**传输控制协议**。一种面向连接的、可靠的、基于字节流的**传输层**通信协议，由 IETF 的 RFC 793 定义。

TCP/IP 协议栈 (TCP/IP protocol stack)

广泛使用的一种协议栈，由从下至上的**物理层**、**数据链路层**、**网络层**、**传输层**、**应用层**组成。

同构连接 (homogeneously connected)

同构网络中的主机被认为是同构连接的。

同构网络 (homogeneous network)

节点和边类型都只有一种的**网络**²⁶。

通信操作

对**数据的操作**的一种。典型的通信操作包括将数据从一个地方传递到另一个地方。

²⁶ 内容来自助教陈国凯推送的网页 https://blog.csdn.net/qq_37430374/article/details/107325554.

统一资源定位 (uniform resource locator, URL)

即 URL。

通用性 (universality)

①抽象三性质之一。计算抽象强调用一个通用抽象代表多个具体需求。它意味着用统一的一套方法处理该层次所有的计算过程，解决该层次的特有问题。这也被称为抽象的泛化能力。

②计算思维的内部特征之一，指计算机能够求解任意可计算问题。

图灵测试 (Turing test)

回答“计算机与人能否被区别开？”这个问题的客观实验方法。一个询问者分别询问一个人和一台计算机，并让询问者指出哪一个被询问者是计算机。如果计算机能够与提问者问答对话，使得提问者误判其为人，那么就说计算机通过了图灵测试。

图灵机 (Turing machine)

图灵机是一个七元组 $\{Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{reject}\}$ ，其中 Q, Σ, Γ 都是有限集合。

- 状态集合： Q
- 输入字母表： Σ
- 带字母表： Γ
- 转移函数： $\delta: Q \times \Sigma \rightarrow Q \times \Gamma \times \{\rightarrow, \leftarrow\}$
- 起始状态： $q_0 \in Q$

· 接受状态: q_{accept}

· 拒绝状态: q_{reject}

单带图灵机是这样完成任务的: 给定一个写着输入的右端无限长纸带, 图灵机从纸带的最左端出发, 按照转移函数进行状态转移以及左右移动和写操作, 最终进入接受状态或拒绝状态。

图灵机可以用**状态转移图**来描述。

图灵问题 (Turing' s problem)

如何把计算机变得更加智能? 更确切地, 如何把**计算系统**变得更加智能? 如果不能, 计算机能够逼近人的智能吗? 这些问题统称为图灵问题。

值得指出, 图灵问题的提法似乎与**无为而治原理**是相背的。徐志伟老师指出²⁷: 计算系统包括计算系统与计算机应用系统, 两个系统都可以做得更加智能。无为而治原理是指导思想, 而不是教条, 且业界的场景、计算机系统的外延在不断进化, “simple”的内涵也有了很大的改变。例如, 提供深度学习功能的加速器 (TPU、GPU、寒武纪等), 现在已经开始成为计算机的标配了, 因为它们已经有 TensorFlow 框架的支撑, 变得足够 simple。

研究图灵问题的领域称为**人工智能或机器智能**。

图形用户界面模式 (graphic user interface mode, GUI mode)

按人机交互方式分类的一种**使用模式**。早期的计算机仅仅支持字符界面, 现在很多计算机都支持**图形用户界面**了。

²⁷ 编者发现了这个矛盾之处, 与徐志伟老师进行了通讯, 他给出了很详细的回复。

URL (Uniform Resource Locator, URL)

即**统一资源定位**，也称**网络地址**，一串指明**网络资源**在网络中的地址的字符串，便于**浏览器**访问该资源。

一个基本的 URL 包括三个部分：

第一部分指明体系，标明一种协议；

第二部分指明主机，标明一个互联网主机的名字；

第三部分指明路径，标明资源存储在主机上的位置。

VPN (virtual private network)

即**虚拟专用网络**。在公用网络上建立安全的专用网络，进行加密通讯。

网关 (gateway)

一种**组网设备**。一个子网络的网关是连接子网络和一个更大的网络的**路由器**²⁸。

网络 (network)

有时也叫**网络计算系统**。很多计算过程需要将多个部件连接在一起形成一个计算系统，这些部件往往被称为**节点**，它可以是抽象或真实的实体²⁹。

网络是节点连接或通信的整体³⁰。网络的节点不必需要和其他节点通信，也可以有**孤立节点**³¹。

²⁸ 这一论断来自英文教材。

²⁹ 这一论断来自大作业 5。

³⁰ 这一论断来自计科导 PPT。

³¹ 这一论断来自大作业 5。

一个计算系统就是由多个节点连接通信而形成的网络。

网络可以是一个硬件系统、软件系统、数据系统、应用服务系统、社会网络系统。例如，全球计算机科学文献网络通过“引用”连接起来，但它们之间并不相互通信。

网络层 (network layer)

TCP/IP 协议栈的第三层，它只有 IP 一种协议，在两台互联网主机间发送一种叫做**数据报**的数据包，但并不保证传递可靠，这种方式也称“尽力交付”。连接的两台互联网主机不必同构连接，但可以被一台或多台路由器相连。

网络地址 (web address)

URL 的别称。

网络攻击 (cyber attack)

网络攻击有多种形式：

- (1) 人；
- (2) 恶意软件；
- (3) 利用硬件的攻击；
- (4) Dos 攻击；
- (4) 垃圾邮件；
- (5) 钓鱼。

网络空间之妙³² (wonder of cyberspace)

计算机科学发展历史上的三大奇妙之一。网络空间使得设计者可以创造物理世界中本不可行的虚拟事物和虚拟世界。计算机科学创造了一种不同于物理空间和人

³² 汉译来自中文课本。

类社会的新空间：网络空间，它由运行在计算机上的程序组成。

网络思维 (network thinking)

强调计算过程中的**连通性**与**消息传递**特征的思维方式。

当代计算机科学发展了**名字空间**和**网络拓扑**概念来体现连通性，发展了**协议栈**技术以实现消息传递。网络思维是名字空间、网络拓扑和协议栈形成的整体思维。网络思维并不必涉及消息传递（或通信协议）³³。

网络拓扑 (network topology)

亦称**网络拓扑结构**，有时简称**拓扑**。它说明节点间可能的连接和连接的实际使用。

任何一个网络，在特定时刻都可以被看成一个由**节点**和**边**构成的**图**。

网络思维关注网络的拓扑及其演变的规律。

根据网络的拓扑随时间变化的情况，经常遇到的网络可分为三类：**静态网络**、**动态网络**、**演化网络**。它们之间可以相互转化。静态网络的某些节点可以用作交换机，从而使一个静态网络变成动态网络。在某一个时刻，当交换机的连通仲裁动作已经确定，一个动态网络就是一个静态网络。甚至演化网络在某一个时刻也可被看作静态网络加以研究。

使用第二代搜索引擎得到的结果比第一代搜索引擎好得多，因为第一代搜索引擎仅利用网页网络的节点，而第二代搜索引擎同时利用节点和边，更好地实践了网络思维。³⁴

³³ 论断来自计科导 PPT。

³⁴ 这一论断来自大作业 5。

网络效应 (network effect)

网络整体价值大于节点价值之和。

描述这一效应的定律有梅特卡夫定律、里德定律。

连接到互联网的笔记本电脑比未联网的笔记本电脑对用户更有价值，因为它受益于网络效应。

网络诱骗 (phishing)

即钓鱼。

谓词 (predicate)

见“谓词逻辑”。

谓词逻辑 (predicative logic)

亦称一阶逻辑。命题逻辑把简单命题作为最基本的单元，不再往下分析。比如说命题“ π 是无理数”和命题“无理数是实数”这两个命题，在命题逻辑的范畴内是找不到什么联系的。谓词逻辑继续拆分命题，把命题拆为“ π ”、“...是无理数”、“...是实数”这些结构，可以得出命题“ π 是实数”这种命题。其中“...是无理数”、“...是实数”称为谓词。

稳定匹配问题

考虑如下场景。

假设 n 名男生和 n 名女生一起参加一个舞会，每一个人都希望找到合适的舞伴。对于每一名女生，根据自己的标准，对参会男生有一个排序。同样，每一名男生对参会女生也有一个排序。假设舞会开始的时候他们任意组成了 n 对舞伴，开始跳第一支舞。

在这个过程中，如果存在着一对男生和女生，他们彼此不是对方的舞伴，但是他们每一个人都觉得对方比自己当前的舞伴更好，那么在下一支曲子开始的时候，他们就会选择对方作为自己的舞伴，更换掉当前的舞伴。我们称这样的一对男女生为一个**不稳定对**。如果一个匹配中存在着一对不稳定对，则称这样的一个匹配是**不稳定的**。反之，称这样的匹配是**稳定的**。

现在的问题是，这 n 名女生与这 n 名男生是否可以一起形成 n 对稳定的舞伴？

物理层 (physical layer)

TCP/IP 协议栈的第一层，通过电缆、光纤或无线电波提供物理通信渠道。它通过发射由 0 和 1 组成的物理信号工作。

无缝衔接 (seamless execution)

亦称**无缝级联**，系统思维的关键概念之一（也是**计算思维的内部特征之一**），即让计算过程在全系统中流畅地运行，不出现或少出现缝隙和瓶颈。

理解无缝衔接需要理解四条原理：**扬雄周期原理**、**博斯特尔健壮性原理**、**冯·诺依曼穷举原理**、**阿姆达尔定律**。前三者主要应对缝隙问题，后者主要应对瓶颈问题。它们合起来使得计算步骤可以级联起来，无缝流畅地实现计算过程。

物理隔离 (physical isolation)

应对网络攻击和网络犯罪的方法之一。切断计算机与互联网的连接。

例如，一些金融机构核心的计算系统便不连接互联网。

无损压缩 (lossless compression)

一种压缩方式，指在压缩文件大小的同时不损失任何信息。

我们压缩一个程序文件或科学数据资料时，应当采用无损压缩，因为任何一个错误的比特都可能会带来整个文件的失效³⁵。

无为而治原理 (the end-to-end argument, E2E principle)

因特网和计算机系统设计的一个基本体系结构原理。计算机系统平台应该只提供必须的、中性的共性技术，尽量让应用去做事，以最好的方式提供应用价值。这条原理和 **KISS 原理** 异曲同工。

正是因为简单、傻瓜，这样的计算机网络才会有中性、灵活的特征。

析取范式 (disjunctive normal form, DNF)

布尔函数的一种范式。

$$f(x_1, \dots, x_n) = Q_1 \vee \dots \vee Q_m$$

其中， $Q_i = l_1 \wedge \dots \wedge l_n$, $l_j = x_j$ 或 $\neg x_j$.

系统架构模型

³⁵ 论断来自英文教材。

人们将“系统由多个模块组合而成”的成功的特定方法称为**系统架构模型**，并赋予特定的名称，如布尔函数的**组合电路模型**、时序电路的**自动机模型**、计算机硬件系统的**存储程序计算机模型**（也称**冯·诺依曼模型**）、数据管理系统的**关系数据库模型**，等等。

系统架构三问题

模块化方法需要回答的三个问题。

- (1) 系统是由哪些模块组成的？也可以反过来问，一个系统如何分解成多个模块？
- (2) 系统是由这些模块如何组成的（模块之间如何连接、有什么接口）？
- (3) 计算过程在系统中如何执行？

系统软件 (system software)

软件栈的组成部分，基础软件的组成部分。包括**语言、编译器、解释器，操作系统和固件**三层。固件可以看成是最底层、最简单的系统软件。

例如，Go 编译器是系统软件；Go shell、Linux shell 和 Python shell 也是系统软件，因为它们都是解释器；BIOS 属于是固件，它是一组固化到计算机内主板上一个 ROM 芯片上的程序，因而也是系统软件。

系统思维 (systems thinking)

见“计算系统思维”。

线路交换 (circuit switching)

也称**电路交换**。线路交换的基本工作原理是：在数据传输期间，源结点与目的结点之间有一条由中间结点构成的专用物理连接线路，在数据传输结束之前，一直保持这条线路。利用线路交换进行通信需以下三个阶段：(1) 线路建立；(2) 数据传输 (3) 线路拆除。其优点是使用专用信道，数据传输迅速、可靠、不会丢失、有序。其缺点是双方之间暂时没有数据传输时，造成带宽资源浪费。因此，电路交换适用于数据传输量大、可靠性要求较高的情况³⁶。

线上线下服务 (online to offline service, O2O service)

融合了计算的虚拟世界与现实世界的服务。

消费者计算

主要用户是消费者个人或家庭的计算，应用类别多，涵盖生活、学习、工作、娱乐等。

写回

指令流水线四个操作阶段之一。将运算结果写回到**寄存器堆**里，或使用存储器地址实现**取数**或**存数**操作。

协议 (protocol)

网络思维的一个核心抽象。确定节点集合以及两个或多个节点之间连接与通信的

³⁶ 参考百度百科。

规则。协议的一个基本要求是无歧义地、足够精确地描述网络连接与通信的操作序列，而且每个基本动作应该是可行的。此外，还需要有助于解决资源冲突、异常处理、故障容错等问题。

协议栈 (protocol stack)

计算思维的内部特征之一，网络思维的重要概念。

在一个使用的网络中，一个协议往往不够，需要几个相互配合的协议一起工作，这些相互配合的协议按层分布，称为一个协议栈。

一个常用而典型的例子是 HTTP/TCP/IP 协议栈。

信息技术 (information technology)

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特指“计算机科学技术”，并采用一个更大的术语信息和通信技术指代“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：硬件产品、软件产品、信息服务。

信息技术产业提供计算机软件产品、计算机硬件产品、网络硬件产品和互联网服务³⁷。

据 2019 年的数据估计，世界范围内平均 1000 个人中才有 1 个信息技术专家，即占 0.1%的比例。

信息通信技术 (information and communication technology, ICT)

³⁷ 论断来自大作业。

国际上往往采用**信息技术**特指“**计算机科学技术**”，并采用一个更大的术语**信息和通信技术**指代“**计算机科学技术与通信技术**”。信息通信技术产业除了信息技术产业可以为我们提供的产品外，还可以为我们提供**电信服务**。

2019 年，世界范围内的信息通信技术花销为 4 万亿美元，然而根据布唐蜜蜂譬喻，世界数字经济的总价值要比它还高，2016 年估计为 15 万亿美元。

信息隐藏原理

在设计或理解一个计算系统时，不论采用或提出什么体系结构模型（系统架构模型），都可以利用一个重要的原理来控制系统的复杂度，即**信息隐藏原理**。

它有以下三个要点：

(1) **隐藏内部信息**：每个模块仅仅暴露其**接口**，以及通过接口可见的外部行为，隐藏该模块的所有内部细节行为和所有信息。

(2) **区分规范与实现**：在设计或理解一个系统时，精确地给出每个模块的**规范**，即其接口和外部行为规定，独立于该模块的**内部实现**。任何符合规范的实现都可以使用。

(3) **抽象并重用模块**：给出每种模块的抽象化表述，给予每种模块抽象特定的命名指称，并尽量重用模块的抽象。

虚拟专用网络 (virtual private network, VPN)

即 VPN。

延迟 (latency)

传输信息所花费的时间。

演化网络 (evolutionary network)

按网络拓扑变化分类的一种网络。节点和边都随着时间改变的网络。

例如，万维网上所有网页构成的网络便是一个演化网络。

扬雄周期原理 (Yang' s cycle principle)

理解无缝衔接需要的原理之一。《太玄经·周首》中有言：“阳气周神而反乎始，物继其汇。”一个计算过程是由若干个步骤或子过程组成的。一个子过程执行并完成了自身的功用，系统就回到某个起始状态，重新执行下一个子过程。万物在过程的展开中各自呈现自己的特色。

异常 (exception)

常见的三类异常是中断、硬件出错及保底异常。

移动互联网计算模式 (mobile Internet mode)

按人机交互方式分类的一种使用模式。通过智能手机以及后端的移动互联网，打开微信等应用的模式。

异或 (exclusive or)

布尔逻辑连接符 \oplus ， $x \oplus y = 1$ 当且仅当 $x \neq y$ 。

异或门 (XOR gate)

一种门。



一阶逻辑 (first-order logic)

即谓词逻辑。

译码 (instruction decode, ID)

指令流水线四个操作阶段之一。控制器解析当前指令，产生控制信号。

易失性存储器 (volatile memory)

断电后内容丢失的存储器。通常比非易失性存储器 (NVM) 要快。有两种常见的类型：动态随机存取存储器 (DRAM) 和静态随机存取存储器 (SRAM)。

硬件错误 (hardware error)

一种异常。当计算机的内存硬件出错时，计算机系统会立即进入一个事先设计好的异常处理程序，不使用内存，处理此类异常。

应用层 (application layer)

TCP/IP 协议的第五层。由于可能有很多层应用堆叠，我们只考虑使用 HTTP 协议的网络应用。它的目的是获取超链接的超文本资源。

有限性

抽象三性质之一。从多个层次（角度，视野）理解一个系统，每个层次仅考虑有限的、该层次特有的问题，忽略其他层次，忽略同一层次的其他问题。


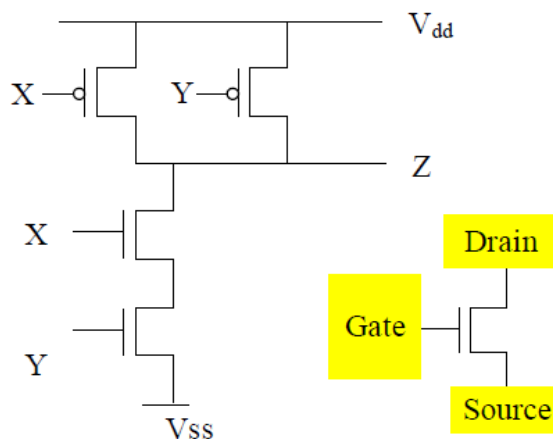
语法错误 (syntactic error)

指程序的语法有误，编译器无法把它转变为适当的编程语言。语法错误只在编译期间出现，它是编译时错误的一种。

与非门 (NAND gate)

由两类共四个晶体管组成，上面两个负晶体管并联，下面两个正晶体管串联，每个输入变量同时接到一个负晶体管和一个正晶体管。

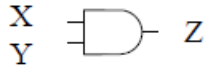
X	Y	Z
0	0	1
0	1	1
1	0	1
1	1	0

当任何一个输入变量为低电平（逻辑 0）时，相应的负晶体管导通，同时相应的正晶体管断开，这使得输出变量 Z 连接到高电平 V_{dd} （逻辑 1）。当两个输入变量同时是高电平时，两个正晶体管导通，两个负晶体管断开，相当于有一根导线从低电平 V_{ss} 连接到 Z，即 $Z = 0$ 。

与门 (AND gate)

一种门。



域名系统 (Domain Name System, DNS)

即 DNS。

元数据 (metadata)

关于数据自身的数据 (data about data)，提供关于数据的附加信息，包括数据的格式和组织方式、文件名、文件大小、运行许可、创建时间等。

包头也称为元数据³⁸。

原子命题 (atomic proposition)

亦称**简单命题**。不包含其他命题作为其组成部分的命题，即在结构上不能再分解出其他命题的命题。它不可以包含**逻辑连接词**。

云计算模式 (cloud computing mode)

一种**互联网计算模式**。将大部分资源放在服务器端 (云端)，客户通过互联网使用，包括硬件资源、软件资源、数据资源。

³⁸ 论断来自计科导 PPT。

运算操作

对数据的操作的一种。典型的运算操作包括加、减、乘、除、与、或、非等算术逻辑运算。

运算器

ALU 的别称。

运行时错误 (run-time error)

程序编译完成后，执行时出现的错误。

真值表 (truth table)

可以将所有变量的全部可能取值以及所对应的布尔函数值都列在一个表里，此即一布尔函数的真值表。

征求意见稿 (request for comments, RFC)

ARPANET 的研制带来的新概念。开放的技术标准文档，大部分 RFC 直接作为技术标准使用。

RFC 文档是一系列关于 Internet (早期为 ARPANET) 的技术资料汇编。这些文档详细讨论了计算机网络的方方面面，重点在网络协议，进程，程序，概念以及一些会议纪要，意见，各种观点等。RFC 享有“网络知识圣经”之美誉³⁹。

³⁹ 论述来自百度百科。

世界上第一个 RFC 是加州大学洛杉矶分校的一名研究生斯蒂夫·克罗克尔写的，定义了网络节点间的软件协议。

正确性 (correctness)

计算思维的内部特征之一。计算机求解问题的正确性往往可以精确地定义并分析。

指令 (instruction)

计算机硬件能够理解和执行的基本步骤。

指令集 (instruction set)

指令集抽象的组成部分。一台计算机可执行的所有指令的集合，它组成提供给软件的指令接口。软件按照指令集中的指令去编写，即可被该计算机正确地处理⁴⁰。

指令集抽象 (instruction set abstraction)

世界上有无穷个计算过程，如何能够考虑到它们的所有步骤的需求？计算机科技工作者采用指令集抽象来解决这个问题，包含下面三个要点：

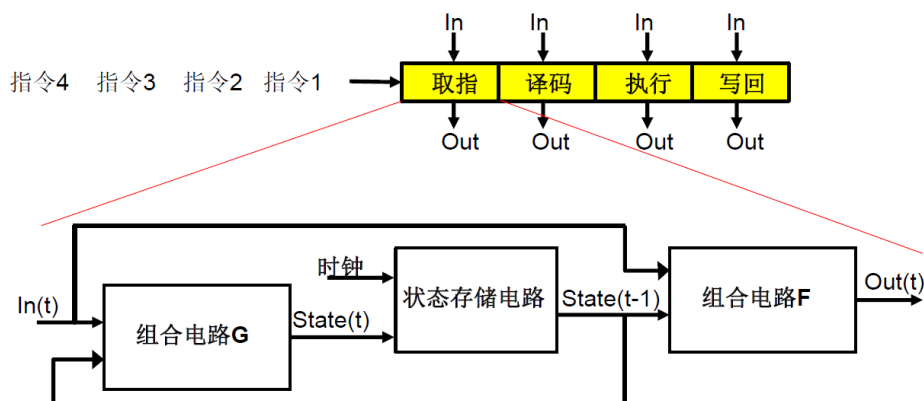
- (1) 指令集；
- (2) 指令流水线；
- (3) 下一指令：提供一种机制，确保一条指令执行完毕后，能够确定并自动执行下一条指令。今天的计算机将当前正在执行的指令放在指令寄存器中，下一条指

⁴⁰ 论断来自陈国凯。

令的地址放在**程序计数器**中解决这个问题。

指令流水线 (instruction pipeline) ⁴¹

指令集抽象的组成部分，当代**处理器**的核心⁴²。每条指令的执行都可以被映射到自动机的一次变换或多次变换的序列。这些变换构成了指令流水线的**阶段**。本课程中的指令流水线包含**取指、译码、执行、写回**四个操作阶段，这四个操作都对应于自动机的一次变换。为了提高速度，多个指令的这四个操作可以通过一个指令流水线的方式重叠执行⁴³。



任何程序、所有指令都是由指令流水线执行。指令流水线是**时钟同步**的时序电路。

指令驱动 (instruction driven)

冯·诺依曼体系结构的要点之一。计算机仅当一条指令被执行时改变它的状态。这里的“状态”指**存储器**和**寄存器**的内容。

指令周期 (instruction cycle)

⁴¹ 徐志伟老师很喜欢考这个，大家要会画指令流水线的这两幅图哦。

⁴² 论断来自计科导 PPT。

⁴³ 此处的“重叠执行”指不同指令的执行在时间上有重叠。

一条指令从时间上的开始到结束构成一个指令周期，它事实上是通过多个**时钟周期**构成的，其执行体现为周而复始地执行一条又一条时钟周期的过程。

执行 (instruction execute,EX)

指令流水线四个操作阶段之一。处理器执行当前指令的操作码规定的操作，如算术逻辑运算，或计算出访问指令所需要的存储器地址。

指令寄存器 (instruction register,IR)

处理器组成部分之一。用于存放正在被执行的指令。

智能的 (smart/intelligent)

从计算机科学的本源看，今天说一个系统是智能的，大体上有三类含义：

- (1) 计算机化。系统已经具备计算机的基本特点，即自动执行程序，并且可编程、可演化，不是一个固定的、僵硬的系统。
- (2) 密集计算。系统执行大量复杂计算，产生非平凡的结果，而不是简单地采集数据并对原始数据做简单处理。
- (3) 人工智能。系统具备传统人工智能特征，如呈现感知计算与认知计算能力。

指数退避 (exponential backoff)

多路访问中决定发生冲突后随机等待时间的一种算法，一般采用的是**二进制指数退避**。

指数之妙⁴⁴ (wonder of exponentiation)

计算机科学发展历史上的三大奇妙之一，指计算资源随时间指数增长，这种过程已经持续了数十年，且很有可能会持续下去。

有三个定律可以描述指数奇观：**诺德豪斯定律**、**摩尔定律**、**科克定律**。

指数诅咒 (curse of exponentiation)

指数奇观的对立面。一些问题和算法的复杂度是指数级的，但这些困难也促成了很多有趣的研究和创新。

指针 (pointer)

指针也就是内存地址，指针变量是用来存放内存地址的变量。以 C 语言为例，每一个变量都有一个内存位置，每一个内存位置都定义了可使用&运算符访问的地址，它表示了在内存中的一个地址。使用一元运算符*来返回位于操作数所指定地址的变量的值。

中断⁴⁵ (interrupt)

一种**异常**。计算机运行过程中，出现某些意外情况需主机干预时，机器能自动停止正在运行的程序并转入处理新情况的程序，处理完毕后又返回原被暂停的程序继续运行。

⁴⁴ 汉译来自中文课本。

⁴⁵ 这个词条摘编自百度百科，因为课本上没有严格定义。

中间件 (middleware)

包括数据库软件、浏览器软件和万维网服务器软件等。之所以被称为中间件，是因为它们在系统软件之上、执行软件之下。虚拟机也可以被看成一种中间件。

中央处理器 (central processing unit, CPU)

见“处理器”。

主定理 (master theorem)

一个经常被用于分析许多重要分治算法的复杂度的定理。

设 f 是满足递推关系

$$f(n) = af\left(\frac{n}{b}\right) + cn^d$$

的增函数，其中 $n = b^k$, k 是一个正整数， $n \geq 1$, b 是大于 1 的整数， c, d 是实数，满足 $c > 0, d \geq 0$ 。那么

$$f(n) = \begin{cases} O(n^d), & a < b^d \\ O(n^d \log n), & a = b^d \\ O(n^{\log_b a}), & a > b^d \end{cases}$$

主机 (host)

网络术语。任何连接到网络的设备都称为主机。

直接被用户使用的主机叫做客户端，其发送的请求被另一台主机响应，称为服务器。须注意主机不包括组网设备。

专门用途寄存器 (special-purpose register)

见“寄存器”。

转义字符 (escape character)

亦称**换码符**。所有的 **ASCII** 码都可以用“\”加数字来表示。而 C 语言中定义了一些字母前加“\”来表示常见的那些不能显示的 ASCII 字符，如 \0, \t, \n 等，就称为转义字符，因为后面的字符，都不是它本来的 ASCII 字符意思了⁴⁶。

常用的转义字符及其十进制下的 ASCII 码值为：

\b 退格 (BS, backspace), 008

\n 换行 (LF, Line Feed), 010

\f 换页 (FF, NP form feed, new page), 012

\r 回车 (CR, Carriage Return), 013

状态电路

亦称**存储电路**。能够存储状态的电路。可以通过含有反馈回路的电路（触发器和门门）、静态存储器、动态存储器构造。

一个状态电路可以由一个或多个**存储单元**或**触发器**来实现⁴⁷。

自动执行 (automatic execution)

亦称**机械执行**。同时属于**计算思维**的**内部特征**和**外部特征**。它的含义如下：

①任何计算过程都是在计算系统（也称计算机）中执行的。

⁴⁶ 内容参考自谭浩强。

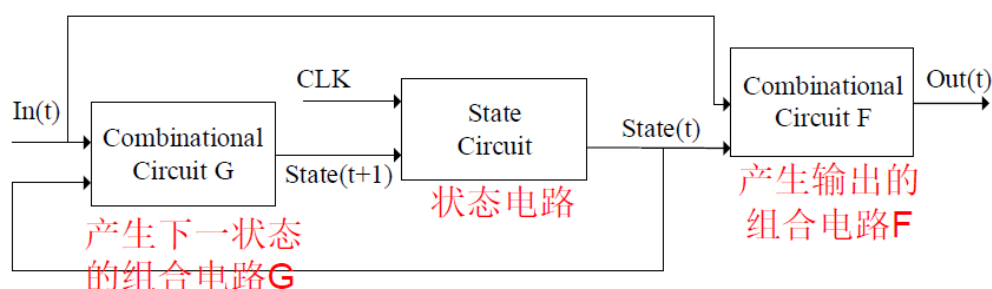
⁴⁷ 这个论断来自大作业 5.

②一个计算过程在计算系统中自动执行，是指从计算过程中的第一个步骤到最后一步的整个执行过程中，计算系统自动地（机械地）执行所有步骤，不需要人工干预。

另一种说法是，计算机能够自动执行由离散步骤组成的计算过程⁴⁸。

自动机 (automaton)

由一个系统时钟驱动的时序电路，也称时钟同步的时序电路。它在时刻 $t=1,2,3,\dots$ 执行一系列变换，产生一系列状态值和输出值。其中，状态电路是一个或多个记忆单元。



带时间的两个布尔函数，即输出函数 F 和状态函数 G ，分别是

$$Out(t) = F(In(t), State(t - 1))$$

$$State(t) = G(In(t), State(t - 1))$$

自动机还可以用状态转移图表示。

字节 (byte)

亦称八位，即八个比特。

⁴⁸ 这种说法来自计科导 PPT，前面的详细说法来自中文教材。

子系统

模块的另一种说法，见“模块化”。

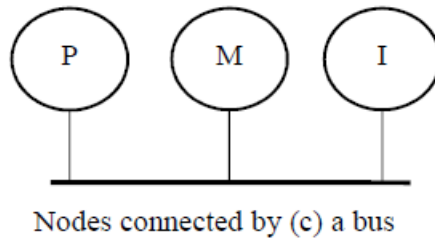
自由软件 (free software)

理查德·斯多曼提出的与**专有软件**相对的概念，它具有四种自由：

- (1) **使用的自由**：用户有权利以任何方式为任何目的使用一个软件；
- (2) **修改的自由**：用户有权利修改软件以满足自己的需求，这意味着用户有权利获取和修改软件源码；
- (3) **分发的自由**：用户有权利复制软件，并把副本分发给其他任何人。分发的副本可以是免费的，也可以是收费的，用户有权利选择免费分发或收费分发的具体方式；
- (4) **分发修改过的软件**的自由：用户有权利以免费或收费方式将修改后软件分发，以便于社会能从软件的改进中受益。

为了保障这四种自由，斯多曼在一些法律教授的帮助下设计了新的软件许可证：**GNU 通用公共许可证**。随着时间推移，该许可证也得到了修改和扩充，出现了十余种自由软件许可证，人们用**开放源码**来泛指这些许可证。

总线 (bus)



组合电路

不含回路的**逻辑电路**。另一种说法是，由与、或、非门电路组合而成的逻辑电路。

⁴⁹可以证明， m 个输入 n 个输出的互不等价的组合电路有 $(2^n)^{2^m}$ 种。

组网设备 (networking device)

包括接入点、集线器、交换机、路由器、网关。

⁴⁹ 前一种说法来自中文教材，后一种说法来自计科导 PPT.