计算机历史

公元前2500年，出现算盘，10进制

1613年出现“computer”一词，早起代表一个职业。在1800年之后逐步替代成机器。

1694年，步进计算器，

计算机先驱Charles Babbage

1823年，提出差分机…

布尔逻辑和逻辑门

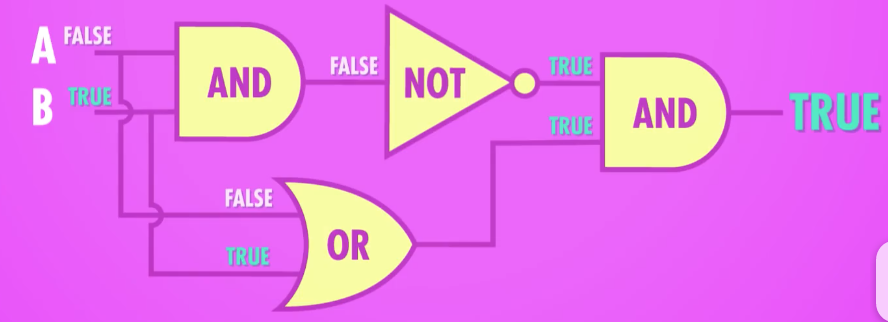
and与：都为”真“，输出”真“；有”假“，输出”假“。

|  |  |  |
| --- | --- | --- |
| inputA | inputB | output |
| true | true | true |
| true | false | false |
| false | true | false |
| false | false | false |

or或：都为”假“，输出“假”；有“真”，输出”真“

|  |  |  |
| --- | --- | --- |
| inputA | inputB | output |
| true | true | true |
| false | true | true |
| true | false | True |
| false | false | false |

Xor 异或：



|  |  |  |
| --- | --- | --- |
| inputA | inputB | output |
| true | true | false |
| false | true | true |
| true | false | True |
| false | false | false |

二进制

1字节=8bit =2^8=256

MB百万字节（million bytes）

GB 十亿字节（giga is billion bytes）

TB （terabyte 8 trillion bytes）

1KB（kilobyte）=2^10=1024bit

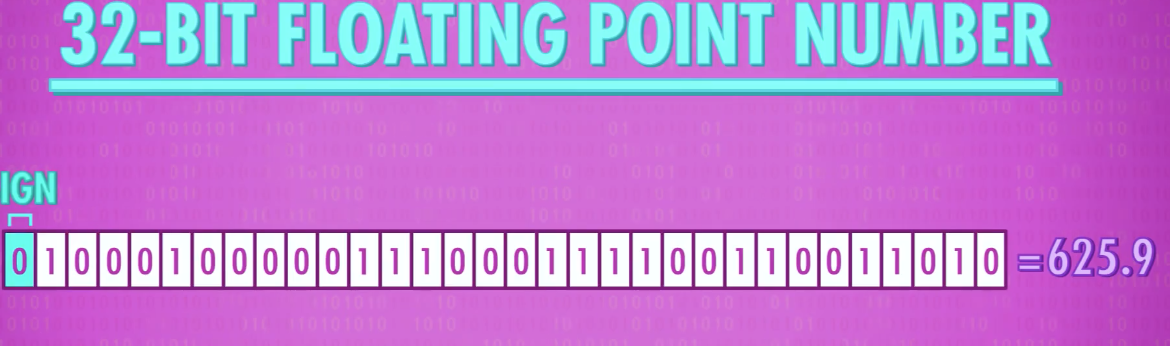
32位float储存

第一位是符号，判断正负；后8位是指数，2^E 表示指数位；最后23位是有效数字。

625.9=1001110001.1110011001100110011001100110011001100110011

10001000=136 ；进位136-127=9小数点向前进9位（E）

.00111000111100110011010为有效数字（M）。





IEEE 754 对有效数字M和指数E，还有一些特别规定。

前面说过，1≤ M <2，也就是说，M 可以写成 1.xxxxxx 的形式，其中 xxxxxx 表示小数部分。IEEE754 规定，在计算机内部保存 M 时，默认这个数的第一位总是 1，因此可以被舍去，只保存后面的 xxxxxx 部分。比如保存 1.01 的时候，只保存 01，等到读取的时候，再把第一位的1加上去。这样做的目的，是节省 1 位有效数字。以32位浮点数为例，留给 M 只有 23 位，将第一位的1舍去以后，等于可以保存 24 位有效数字。

E 的真实值必须再减去一个中间数，对于 8 位的 E，这个中间数是 127；

ascII码，有8位，能储存255个字符。汉字日文有上千个字符所以用2字节，16位。

算数逻辑单元ALU

ALU 有两个单元，1个算数单元，1个逻辑单元。

算数单元，负责计算机里的所有数字操作

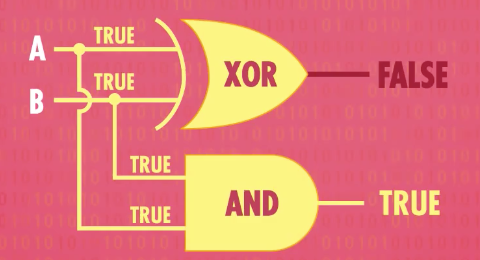
xor用作一位加法器

|  |  |  |
| --- | --- | --- |
| inputA | inputB | output |
| 1 | 0 | 1 |
| 0 | 1 | 1 |
| 0 | 0 | 0 |

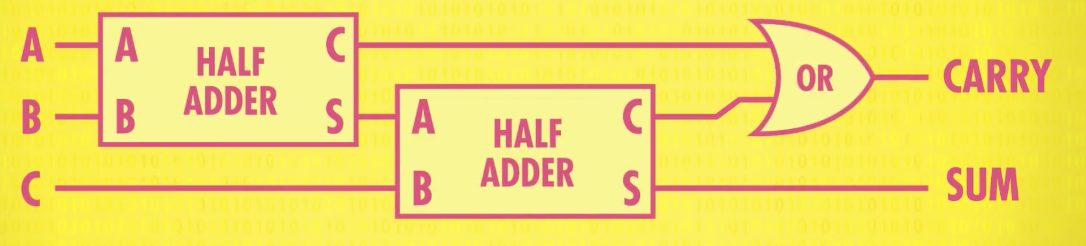
当1+1出现需要进位1bit不够，需要进位，此时需要添加AND门。

半加器出现

 And门处理进位

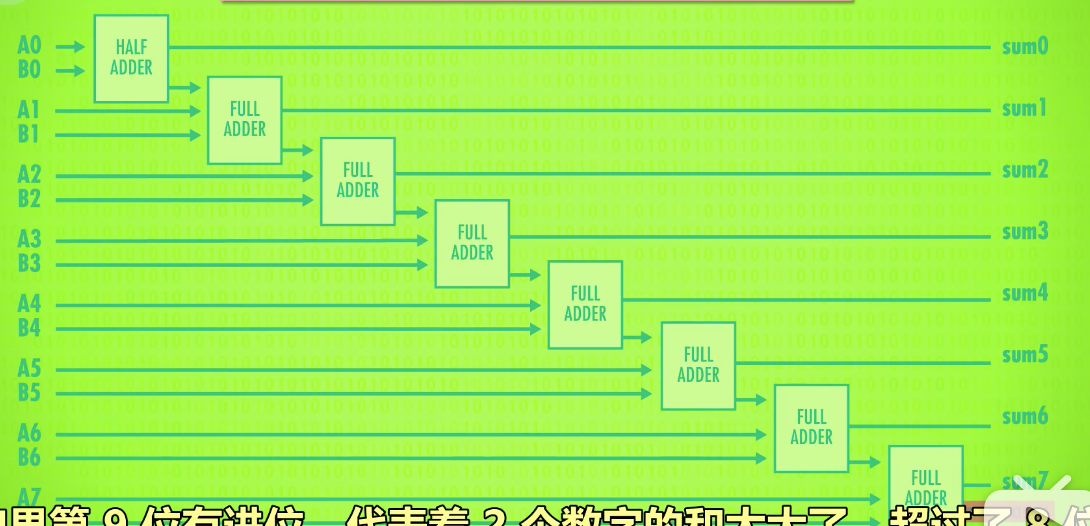


超过1+1的运算，如1+1+1，需要使用全加器



8位计算，继续进位会溢出，需要更多全加器。

半加器进位与下一个全加器相加



逻辑单元

操作码代表A和B执行什么操作。flag代表状态，A-B，flag为0代表AB相等。

negative位true说明A<B.

图片包含 物体

描述已自动生成

寄存器和内存

1个寄存器储存1个数字，数字多少位称为位宽。

# -

从头开始学

第10集:早起的编程方式

1801年，约瑟夫。玛丽。雅卡尔发霉了可编程纺织机。

1890年美国人口普查，用穿孔纸卡，汇总数据。

1940年，IBM 402核算机，负责算盈亏总额。

1946年，世界第一台通用电子计算机，ENIAC，用大量拆下板，程序在纸上设计好后，给ENIAC连线，，将近三周时间，很繁琐。

1950年代，存储程序计算机出现，用内存储存程序，方便cpu读取，程序和数据都存在“冯诺依曼结构”。命名自 约翰。冯。诺依曼

冯诺依曼计算机的标志是，一个处理器+数据寄存器+指令寄存器+指令地址寄存器+内存，由曼彻斯特大学于1948年建造，绰号“宝宝👶”

1980年代，几乎所有计算机都有穿孔打卡器读取数据，还有插线板。

第11集：编程语言的发展史

第一条指令在内存地址0：0010 1110（前4位是操作码，后四位是内存地址） 意思是读取内存地址14，放入LOAD\_A代表的(0010指向的寄存器)。

早起用机器码写程序。pseudo-code(伪代码)对应机器码。

助记符：用指令代替0和1的机器码。通过程序将文字转换成二进制指令，这样cup可识别文字指令。

汇编器读取“汇编语言”写的程序转换成机器码，汇编器能自动分析跳转地址。汇编器不用固定的跳转地址，可以插入可跳转的标签，当程序被传入汇编器，汇编器会自己搞定跳转地址。

编译器:将高级语言转换成低级语言。

第12集编程原理-语句和函数

编程语言让程序员专心解决问题，不用管硬件细节。

第13集算法入门

归并排序：

最短路径：

第14集数据结构

数组：数组是可以再内存中连续存储多个元素的结构，在内存中的分配也是连续的，数组中的元素通过数组下标进行访问，数组下标从0开始

链表链表是物理存储单元上非连续的、非顺序的存储结构，数据元素的逻辑顺序是通过链表的指针地址实现，每个元素包含两个结点，一个是存储元素的数据域 (内存空间)，另一个是指向下一个结点地址的指针域。根据指针的指向，链表能形成不同的结构

栈是一种特殊的线性表，仅能在线性表的一端操作，栈顶允许操作，栈底不允许操作。 栈的特点是：先进后出，或者说是后进先出，从栈顶放入元素的操作叫入栈，取出元素叫出栈。

队列队列与栈一样，也是一种线性表，不同的是，队列可以在一端添加元素，在另一端取出元素，也就是：先进先出

树是一种数据结构，它是由n（n>=1）个有限节点组成一个具有层次关系的集合

图是由结点的有穷集合V和边的集合E组成。其中，为了与树形结构加以区别，在图结构中常常将结点称为顶点，边是顶点的有序偶对，若两个顶点之间存在一条边，就表示这两个顶点具有相邻关系。

第15课 阿兰图灵

图灵完备：

图灵测试：机器能让人类相信他是人类

第16课软件工程

面向对象编程

版本管理

集成开发环境

第17课 集成开发环境

第18课 操作系统

操作系统是程序与机器之间的媒介。

动态内存分配

第20课文件与文件系统

文件在系统里是碎片化储存的，一个文件会占用一个或多个块。在文件目录文件中会显示每个文件占用那些块，同时会储存源数据（文件名，创建时间，修改时间，权限等）

文件的删除只是删除目录里的记录，数据还是存在块中，当下次有文件需要占用这个块，数据才会被替换。

文件的移动，只是

文件的删除只是删除目录里的记录，数据还是存在块中，当下次有文件需要占用这个块，数据才会被覆盖。

文件的移动是将文件信息从一个目录转移到另一个目录。

第21课 压缩

无损压缩：没有丢失信息

游程编码：减少重复信息

dftba：用字典储存代码和数据的关系，霍夫曼树。

有损压缩：音频MP3：用不同精度编码不同频段，人对于低音感知不强，只能感到震动。wav和flac是无损。

图像jpeg：将图片分解成8\*8像素块，然后删除大量高频率像素块。

视频mpeg-4：只存帧与帧之间变化的部分。

第22课 命令行界面

1970年之前，用打字机与机器对话，后来屏幕普及，用屏幕显示计算机反馈的信息，而且便于修改输入错误。

命令行游戏：Zork

现在多用于访问远程服务器

第22课命令行界面

第25课个人计算机革命

解释器和编译器的区别：编译器是提前转换，解释器运行时转换

第26课图形用户界面

1968年恩格尔巴特，在秋季计算机联合会议展示了他的系统，这次演示被视为如今所有演示的祖先。包括，位图图像，文字处理，视屏会议，实时协作编辑文件。商业上失败了。

施乐之星上首次出现“复制，粘贴”，这台计算机上市前，乔布斯参观了实验室对

第28课：计算机网络

计算机近距离构成的小型网络叫局域网（LAN），以太网需要每台计算机有唯一的媒体访问控制地址，即MAC地址。发送数据时，将接收方的MAc地址作为数据的前缀，发送到网络中。

多台计算机共享一个传输媒介，这种方法叫“载波侦听多路访问”，即CSMA。

载体传输速度叫带宽。

载体和其中的设备总称为冲突域。利用交换机将大的冲突域分割成小的冲突域，交换机记录那些MAC地址在那个冲突域里。

路由：在大型网络中，一个地点到另一个地点有多条路径。

专线：这条线路无需共享，自己能随意使用。一般军队，银行和一些机构在使用。

报文交换：可以使用不同的路由传递，使通信更可靠更能容错。

跳数：信息沿着路由的跳转次数。

跳数限制：跳数会记录在信息中心，传输时会更新跳数，检测到某条信息的跳数过高，就知道了路由出现了问题。

数据包：将大报文分成很多小块。数据包中包含目标地址，报文具体格式由互联网协议定义，简称IP。传输大报文的时候会阻塞宽带，之后有小报文传输也需等之前的传输完成，造成宽带阻塞。

TCP/IP:解决数据包乱序问题。同一报文的不同数据包，会经过不同路线到达顺序可能不同。

第29课 互联网

广域网：WAN

互联网提供商：ISP。互联网提供商通过路由器的WAN口提供网络。

请求YouTube是视频时，将请求数据包传输到互联网主干，主干再到达YouTube的服务器。traceroute查看跳数

UDP：简单快捷。根据端口号将数据包送到目标程序，UDP头部有校验和用于检查数据是否正确。然而UDP不提供数据重发和数据修复机制，接收方得知数据损坏后将数据包丢掉，发送方不知道数据是否传输成功。ip负责将数据包送到正确的计算机，UDP将数据包送到正确的程序。

TCP:可靠。tcp头部也有端口号和校验和，数据包有序号，会同时发送多个包。根据确认码的成功率和来回时间判断网络堵塞程度，调整同时发包数量。

三次握手：sender发送数据包到receiver，receriver接受到数据包后，返回ack1（确认码）。没有收到确认码，会重发。

域名系统：DNS，浏览器输入网址后去DNS服务器查找对应的IP，如果存在会返回IP到浏览器。DNS是树状结构，顶级域名TLS（.com.gov.net.org等），二级域名（google.com），三级域名（image.google.com）。

开放式系统通信参考模型（OSI）

数据链路层操作物理层：MAC，碰撞检测，指数退避等底层协议。

网络层：负责各种报文交换和路由。

传输层：UDP和tcp负责计算机间点对点传输。检测和修复错误。

会话层：使用udp和tcp创建连接，传递信息和关闭连接

表示层：

应用层：

第30课 万维网

万维网的基本单位是单个页面。

统一资源定位器：URL，

超文本传输协议：HTTP，

第31课计算机安全

计算机攻击从三个方面进行的：保密性完整新和可用性。

保密性：只有有权限的人才能读取计算机系统和数据

完整性： 使用和修改系统和数据

可用性：有权限的人应该随时可以访问系统和数据

威胁模型分析：对攻击者大致描述，能力，目标，手段（攻击矢量）。

两个主要的安全问题：1. 你是谁？

2. 你能访问什么？

身份认证：生物识别，双重或多重因素认证。

访问控制：通过权限和访问控制列表（ACL）来实现。其中描述了用户对每个文件，文件夹和程序的访问权限。

第32课 黑科&攻击

社会工程学：通过欺骗别让别人泄密。网络钓鱼。

物理接触计算机连接几根线在内存上复制信息。

远程攻击：利用漏洞获取莫衷权限或能力。缓冲区溢出。

代码注入：sql，select或drop语句。

蠕虫：利用漏洞让病毒在电脑间传播，黑科拿下大量电脑后利用这些电脑主城僵尸网络。

拒绝服务攻击：DDoS，发送大量垃圾信息阻塞网络。

第33课 加密

替换加密：把字母变替换成其他字母。

Aes：128位密钥

密钥交换：俩个人将自己的密钥与公共密钥相加再交换，再将交换后的密钥加上自己的密钥，如果两方匹配，配对成功。

迪菲赫尔曼密钥交换：在diffie-Hellman中，单项函数是模幂运算。首先，有个公开的值，基数（B）和模数（M），为了安全想对付发消息，选秘密指数为X，然后算B^XmodM给对方。对方选秘密指数为Y，将B^XmodM的结果给我。公共密钥，B^XYmodM。

对称加密：双方用一样的加密解密信息。

非对称加密：公钥加密信息，只有有私钥的人能解密。反之亦然。多用于签名加密。目前流行的有：RSA。

第34课 机器学习和人工智能

机器学习：统计学算法

分类算法叫分类器，特征数据：将图片，音乐等数据化，标记数据：给分类命名，决策边界：给分类设定边界。

记录正确数和错误数的表叫混淆矩阵。

机器学习算法的目的是最大化正确分类和最小化错误分类。

决策树：把决策空间分成

支持向量机

人工神经网络：ANN

深度学习：

弱AI：只能做特定的事务。

强AI：像人一样聪明的AI。

强化学习：学习什么管用，什么不管用，自己发现成功的策略。

第35课 计算机视觉

RGB值：通过组合红绿蓝三种颜色的强度能够得到任何颜色。

颜色跟踪算法：通过颜色识别某一样物体：先选中物体的特征颜色，在屏幕上扫描每个像素，框出其中颜色最近特征颜色的物体。

边缘检测：将图片的每个像素点的八个方向的灰度与核相乘，再相加，得到的值替换当前像素点。最后形成的图像会增强边缘的显示。

核：判断垂直边缘

|  |  |  |
| --- | --- | --- |
| -1 | 0 | 1 |
| -1 | 0 | 1 |
| -1 | 0 | 1 |

判断水平边缘

|  |  |  |
| --- | --- | --- |
| -1 | -1 | -1 |
| 0 | 0 | 0 |
| 1 | 1 | 1 |

这两个边缘增强的核叫prewitt算子。

卷积神经网络：神经网络最基本的单元是神经元。有多个输入，把每个输入乘一个权重值然后求总和。用一堆神经元处理图像数据，每个都会输出一个新图像，本质是被不同核处理了。然后再通过多次卷积，每次都会在上一次基础上增加复杂度，最后输出识别结果。

第36课 自然语言处理 natural language processing

第39课 教育科技

第40课 奇点 天网 计算机未来

奇点：智能科技的失控性发展。

微软联合创始人 保罗艾伦：提出复杂度刹车，计算机发展曲线有可能是s型，随着复杂度增加，进步会越来越难。

非重复性思维型工作，在未来依然不会被计算机替代。如教师，医生，艺术家和科学家。

未来：加密货币，无线通讯3d打印生物信息量子计算。