



复试笔记

问题:

1.嵌入式和普通pc的区别:

嵌入式系统一般指非pc系统,有计算机功能但不能被称之为计算机的设备或器材.它是以应用为中心,软硬件可裁剪的,适应应用系统对功能 `可靠性 `成本 `提及 `功耗等综合性严格要求的专用计算机系统.简单地说,嵌入式系统集系统的应用软件与硬件于一体,类似于PC中的bios的工作方式,具有软件代码小 `高度自动化 `响应速度快等特点,特别适合于要求实时和多任务的体系.嵌入式系统主要由嵌入式处理器 `相关支撑硬件 `嵌入式操作系统及应用软件系统组成.它可是独立工作的器件.个人pc一词源于1978年IBM的第一部桌上型计算机型号PC,在此之前有apple II的个人计算机.能独立运行 `完成特定功能的个人计算机.个人计算机不需要共享其他计算机的处理 `磁盘和打印机等资源也可以独立工作.

2.数据库系统和文件系统的区别:

区别:

- 文件系统用文件将数据长期保存在外存上,数据库系统用数据库统一存储数据.
- 文件系统程序和数据库系统有一定的联系,数据库系统中的程序和数据分离.
- 文件系统用操作系统中的存取方法对数据进行管理,数据库系统用DBMS统一管理和控制数据.
- 文件系统实现以文件为单位的数据共享,数据库系统实现以记录和字段为单位的数据共享,

联系:

- 均为数据组织的管理技术
- 均为数据管理软件管理数据,程序与数据之间用存取方法进行转换
- 数据库系统是在文件系统的基础上发展而来的

3.大数据的特征是什么?

大数据,指无法在一定时间范围内用常规软件工具进行捕捉 `管理和处理的数据集合,是需要新处理模式才能具有更强的决策力 `洞察发现力和流程优化能力的海量 `高增长率和多样化的信息资产

- 特征:
 - i. 容量(Vloume):数据的大小决定所考虑的数据的价值和潜在的信息

- ii. 种类(Variety):数据类型的多样性
- iii. 速度(Velocity):指获得数据的速度
- iv. 可变性(Variability):妨碍了处理的有效地管理数据的过程
- v. 真实性(Veracity):数据的质量
- vi. 复杂性(Complexity):数据量巨大,来源多渠道

4.数据安全是什么?

- **1. 完整性:**

指信息在传输 ` 交换 ` 存储和处理的过程保持非修改 ` 非破坏和非丢失的特征,即保持信息的原样性,使信息能正确生成 ` 存储 ` 传输,这是最基本的安全特征

- **2. 保密性:**

指信息按给定要求不泄漏给非授权的个人 ` 实体或过程,或提供其利用的特性,即杜绝有用信息泄漏给非授权个人或实体,强调有用信息只被授权对象使用的特征.

- **3. 可用性:**

指网络信息可被授权实体正确访问,并按照要求能正常使用或在非正常情况下能恢复使用的特征,即在系统运行时能正确存取所需信息,当系统遭受攻击或者破坏时,能迅速恢复并能投入使用.可用性是衡量网络信息系统面向用户的一种安全性能.

- **4. 不可否认性:**

指通信双方在信息交互过程中,确信参与者本身,以及参与者所提供的信息的真实同一性,即所有参与者都不可能否认或抵赖本人的真实身份,以及提供新的的原样性和完成的操作与承诺.

- **5. 可控性:**

指对流通在网络系统中的信息传播及具体内容能够实现有效控制特性

5.嵌入式系统

嵌入式系统,是一种“完全嵌入受控器件内部,为特定应用而这几的专用计算机系统”,根据英国电气工程师协会的定义,嵌入式系统为控制 ` 监视或辅助设备 ` 及其或用于工厂运作的设备.与个人计算机这样的通用计算机系统不同,嵌入式系统通常执行的是带有特定要求的预先定义的任务.由于嵌入式系统只针对一项特殊的任务,设计人员能够对它进行优化减小尺寸降低成本.嵌入式系统通常进行大量生产,所以单个的成本节约,能够随着产量进行成百上千的放大.

嵌入式系统的核心是由一个或几个预先编程好以用来执行少数几项任务的微处理器或者单片机组成.与通用计算机能够运行用户选择的软件不同,嵌入式系统上的软件通常是暂时不变的;所以经常称为固件.

6.中断在存储中的应用

在请求分页存储管理中,当指令的执行所需内容不再内存中时,发生缺页中断;当缺页调入内存

后,应执行“被中断指令”.另外:缺页中断作为中断与其他中断一样要经历如保护cpu环境,分析中断原因,转入缺页中断处理程序进行处理,恢复cpu环境等几个步骤,但缺页中断又是一种特殊的中断,它与一般中断相比,有着明显的区别,主要表现在下面两个方面:(1)缺页中断时在指令执行期间产生和处理中断信号的.(2)一条指令的执行期间,可能产生多次缺页中断

7.三层交换机和路由器在转发数据上的区别

三层交换机采用一次路由多次转发的方式,也就是说对于到达同一目标的数据只需要第一次进行路由查询,以后就直接转发.而路由器是每一个数据包都是独立路由并转发的.因此在局域网内部,三层交换机的速率要比路由器快很多.

8.云计算

是一种基于互联网的相关服务的增加、使用和交付模型,通常涉及通过互联网来提供动态易拓展且经常是虚拟化的资源.云是网络、互联网的一种比喻说法.获取在图中往往用云来表示电信网,后来也用来表示互联网和底层基础设施的抽象.因此,云计算甚至可以让你体验10万亿次的运算能力,拥有这么强大的计算能力可以模拟核爆炸、预测气候变化、解决全球性问题等等.用户通过电脑、笔记本、手机等方式接入数据中心,按自己的需求进行运算.

9.汇编语言跳转

JMP

10.多处理机

由多个独立的处理机组成,每个处理机都能够独立执行自己的程序.作业和任务的并行.用于提高复杂问题的解决速度.

11.虚拟现实VR

是一套由计算机仿真系统创建出来的虚拟世界.通俗讲,就是使用技术手段,让人身临其境,并可以与这个环境进行交互.这套技术主要包括模拟环境、感知、自然技能和传感器各方面,除了计算机图形技术所生成的视觉感知外,还有听力、触觉、嗅觉等感知技术,以及头盔、手套、鞋子等设备.虚拟现实技术的应用领域非常广泛,如军事、医学、教育、娱乐等.

12.人工智能

人工智能是对人的意识、思维的信息过程的模拟.人工智能不是人的智能,但能像人那样思考、也可能超过人的智能.其所使用的技术旨在根据数据和分析赋予计算机能够做出类似人类的判断.该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等.人工智能是研究人类智能活动的规律,构造具有一定智能的人工系统,研究如何让机器去完成以往需要人的智力才能胜任的工作,也就是研究如何应用计算机的软硬件来模拟人类某些智能行为的基本理论、方法和技术.

13.认知计算

认知计算出自IBM人工智能超级计算机“沃森”的称谓,而现在,它更多代表着一种全新的大数据分析方式.随着信息的增加,计算机可在已有的经验的基础上随着时间推移,以学习的、交互的方式,随着数据的进一步增长逐步提高认知的分析行为,就像大脑会自然而然地做事情.

14.量子计算

量子计算是一种利用量子力学原理来传递和处理信息的计算方式.量子计算机是一种利用量子力学原理来传递和处理信息的计算机.量子计算机的基本计算单元是量子比特,而不是传统计算机的比特.量子比特具有叠加态和纠缠态的特性,这使得量子计算机在某些特定的问题上具有传统计算机无法比拟的优势.量子计算机的发展有望在密码学、材料科学、生物科学等领域带来革命性的变革.

15.深度学习

深度学习Deep Learning的概念源于人工神经网络的研究.机器学习研究中的一个全新的领域,其动机在于建立、模拟人脑进行分析学习的神经网络,它模仿人脑的机制来解释数据,例如图像,声音和文本.自2006年以来,机器学习领域,取得了突破性的进展.图灵实验,至少不是那么可望而不可及的了.在技术手段上不仅仅依赖于云计算对大数据的并行处理能力,而且依赖于算法.这个算法就是deep learning.

16.计算机视觉

计算机视觉是一门研究如何使机器“看”的科学,更进一步,是指用摄像机和计算机代替人眼对目标进行识别、跟踪和测量等机器视觉.计算机视觉是一门研究如何使机器“看”的科学.使电脑处理成为更适合人眼观察或柴松给仪器检测的图像.形象地说,就是给计算机安装上眼睛和大脑,让机器能够感知环境与对象.

17.软件重用、软件重构

软件重用,是指在两次或多次不同的软件开发过程中重复使用相同或相似软件元素的过程.软件元素包括程序代码、测试用例、设计文档、设计过程、需要分析文档甚至领域知识.通常,可重用的元素也称作软构件,可重用的软构件越大,重用的力度越大.软件重构,是指在不改变软件外部行为的情况下改善其内部结构的过程,使软件更具可读性、可维护性和可扩展性,通常包括重构程序模块、优化算法、提高性能等方方面面.

18.设计一个迷宫能用到什么算法?

深度优先搜索算法、广度优先搜索算法、A*算法、Dijkstra算法、Bellman-Ford算法、Floyd算法、Prim算法、Kruskal算法等.

19.动态规划

若要解决一个给定问题,我们需要解其不同部分,再合并子问题的解以得出元问题的解.通常许多子问题非常相似,为此动态规划法仅仅解决每个子问题一次,从而减少计算量.一旦某个给定子问题的解已经算出,则将其记忆化存储,以便下次需要同一个子问题解的时候直接查表

20.嵌入式,互联网+是什么,并且举一个具体应用?

嵌入式系统是一种专门的计算机系统,作为装置或设备的一部分.通常,嵌入式系统是一个控制程序存储在ROM中的嵌入式处理器控制版.“互联网+”是创新2.0下的互联网发展的新业态,是知识社会创新2.0推动下的互联网形态演进机器催生的经济社会发展新形态.

21.编译原理和open flow

openflow,一种网上通信协议,属于数据链路层,能够控制网上交换机或路由器的转发平面,借此改变网上数据包所走的网上路径.

22.分布式计算

分布式计算是一种计算方法,和集中式计算是相对的.随着计算技术的发展,有些应用需要非常巨大的计算能力才能完成,如果采用集中式计算,需要耗费相当长的时间来完成.分布式计算,将应用分解成许多小部分,分配给多台计算机进行处理.这样可以节约整体计算时间.

23.数据库管理系统和操作系统的关系

DBMS是database management system数据库管理系统;OS是operating system操作系统.一般DBMS是安装在OS,DBMS的数据管理可以用OS的文件管理,也可以向OS申请空间然后自己管理数据.

24.无线局域网为什么不用碰撞检测而用避免?

主要有以下两个方面的原因:

- (1)无线信号在空间衰减很大,无线网络适配器发射的信号幅度是接受到的信号的几万`几十万`几百万倍.要实现冲突检测,就要从发射信号中检测出比他小很多的接收信号,硬件设计会变的非常复杂,由此会带来设备成本的增加.从另一方面讲,强度差很大的两个信号,即使冲突,问题也不大,信号还是可以被识别,只是弱信号被湮没.
- (2)无线局域网中,存在屏蔽站问题.

25.网络安全功能

- (1)故障管理

故障管理是网络管理中最基本的功能之一,用户都希望有一个可靠的计算机网络,当网络中某个部件出现问题,网络管理员必须迅速找到故障并且排除.

- (2)计费管理

用来记录网络资源的使用,目的是控制和检测网络操作的费用和代价,它对一些公共商用网络很重要

- (3)配置管理

配置管理同样重要,它负责初始化网络并配置网络,以使其提供网络服务.

- (4)性能管理

不言而喻,性能管理估计系统资源的运行情况及通信效率情况.

- (5)安全管理

安全性一直是网络的薄弱环节之一,而用户对网络安全的要求又相当高

- (6)网络管理

网络管理主要包括流量管理,通过各种网络信息的收集、分析和统计,平衡整个网络的负载,合理分配网络流量,提高网络资源的利用率和整个网络的吞吐率.

26.CSMA/CD 和 CSMA/CA

CSMA/CD:带有冲突检测的载波监听多路访问,可以检测冲突.CSMA/CA:带有冲突避免的载波监听多路访问,通过随机等待时间来减少冲突的可能性.

(1)两者的传输觉知不同,CSMA/CD用于总线以太网,CSMA/CA用于无线局域网.

(2)检测方式不同,CSMA/CD通过电缆中电压的变化来检测,当数据发生碰撞时,电缆中的电压就会随着发生变化.而CSMA/CA采用能量检测(ED)、载波检测(CS)和能量载波混合检测三种检测信道空闲的方式.

(3)WLAN中,对于某个节点来说,其刚刚发出的信号强度要远高于来自其他节点的信号强度,也就是说会覆盖掉.

(4)本节点冲突并不意味着在接受节点有冲突.

27.解释代码复用

函数的作用是复用代码、增加代码可读性、降低变成复杂度,而并不能提高代码的执行速度.

28.osi分基层,哪几层

OSI七层模型:应用层、表示层、会话层、传输层、网络层、数据链路层、物理层.

29.解释网络时延是哪几部分组成

网络时延由四部分组成:发送时延、传播时延、处理时延、排队时延.

30.UML是什么?

UML是统一建模语言

31.机器学习是什么?

机器学习是一门多领域交叉学科,涉及概率论`统计学`逼近论`凸分析`算法复杂度理论等多门学科.专门研究计算机怎样模拟人类的学习行为,以获取新的知识和技能,重新组织已有的知识结构使之不断改善自身的性能.

32.木马和病毒的区别

- (1)病毒时当已感染的软件运行时,这些恶行程序向计算机软件添加代码,修改程序的工作方式,从而获得计算机的控制权.
- (2)木马是指未经用户同意进行非授权操作的一种恶意程序.它们可能删除硬盘上的数据,使系统瘫痪,盗取用户资料等.木马程序不能独立侵入计算机,而是要依靠黑客来传播的,他们常常被伪装成“正常”软件进行散播.
- (3)他们最大的区别就是病毒具有感染性,而木马一般不具有感染性,另外,病毒入侵后立刻有感觉,而木马入侵后希望你每有感觉,这样才有利于它“开展工作”.木马虽然属于病毒中的一类,但是要单独的从病毒类型中间剥离出来.独立的称之为“木马”程序.

33.程序与微程序的区别

程序是一系列机器指令的有序集合,用于解决实际问题,有子程序`分支`循环等结构,存放在主存中,可以更新修改.

微程序是一系列微指令的有序集合,微程序设计时将传统的程序设计方法运用到控制逻辑的设计中,因此在微程序中也可以有微子程序`分支`循环等结构;微程序存储在CM中,只能读出不能写入,CM中的所有微程序解释执行整个指令集中的所有机器指令.

34.如何实现CPU与IO的并行

在多道程序系统中,当需要I/O操作时,不需要CPU直接控制I/O操作的完成,可利用通道`DMA及中断机构来直接控制I/O的完成,这时CPU也不是专门空闲等待I/O操作的完成,而是转去另一道程序.

35.控制器主要几种控制方式?

同步控制`异步控制和联合控制.

36.死锁的原因和必要条件是什么?怎么解决死锁?

原因:

- (1)因为系统资源不足.
- (2)进程的推进顺序不合适.
- (3)资源分配不当.

如果系统资源充足,进程的资源请求都能得到满足,死锁出现的可能性就很低,否则就会因为

竞争有限的资源而陷入死锁.

必要条件

- (1)互斥条件:资源不能被共享.
- (2)持有并等待:已有资源的进程在等待新的资源.
- (3)不抢占:资源不能被抢占,只能继续保持.
- (4)循环等待:存在一组进程,每个进程都在等待下一进程持有的资源.

解决方法

- (1)死锁预防:破坏死锁的四个必要条件之一.
- (2)死锁避免:银行家算法.
- (3)死锁检测与解除:检测死锁,然后解除死锁.
- (4)死锁忽略:不解决死锁,让系统自动重启.

37.离散数学握手定理

在无向图中 $G = \langle V, E \rangle$ 中,每个顶点的度数和等于边数的两倍.

在有向图中 $G = \langle V, E \rangle$ 中,每个顶点的入度和等于出边数的和,所有的节点的度数总和等于边数的两倍.

38.操作系统的五大功能

操作系统的五大功能:进程管理 ` 存储管理 ` 文件管理 ` 设备管理 ` 用户接口.

39. HTML是什么?

HTML(超文本标记语言)是一种用于创建网页的标准标记语言 ` 它通过标记来描述网页的结构和内容 ` 主要用于网页的文本 ` 图像和其他媒体元素的布局和格式化 `。

40.类的继承方式

类的继承方式有三种:公有继承,私有继承,保护继承.

41.网络地址变换协议是什么?

网络地址转换协议(Network Address Translation ` NAT)是一种将私有网络地址转换为公共网络地址的技术 ` 以便在私有网络中访问公共网络 `。

42.决策树学习是什么?

决策树学习是一种常见的机器学习方法 ` 通过一棵树状结构进行决策 ` 树的每个内部节点表示一个属性测试 ` 每个分支表示一个测试结果 ` 每个叶子节点表示一个类标签 ` 该方法直观易懂 ` 适用于分类和回归任务 `。

43.指出动态释放相对于静态释放的优点,并举例说明

- (1) 简洁 `漂亮 `不易犯错

动态语言不用程序员关心冗余的类型信息,让其更加关注少量的 `本质的 `简洁的核心逻辑,这使代码量急剧减少,每行的故障数也能够得到有效的控制.从某种意义讲,动态语言试图最大化释放在程序员的生产力和创造力.

- (2) 灵活 `自由

软件设计大师Dave Thomas生动而有趣地将动态类型系统比喻为“Duck Type”:If it like a duck and quacks like a duck,it must be a duck.这种灵活的类型系统使得程序员可以更加自由地进行编程,不用拘泥于类型的限制,可以更加自由地进行编程.

44.什么是管态,什么是目态,与运行状态的关系是什么?

CPU有两个状态,分别是管态和目态.

管态,即操作系统的管理程序运行时的状态,具有较高的特权级别,也称为特权态 `系统态 `内核态或者核心态.当处理器处于管态时,可以执行所有的指令,包括各种特权指令,也可以使用所有的资源,并且具有改变处理器状态的能力.

目态,即用户程序运行时的状态,具有较低的特权级别,又称为普通态或用户态.在这种状态下不能使用特权指令,不能直接使用系统资源,也不能改变CPU的工作状态,并且只能访问这个用户程序自己的存储空间.

用户态不允许程序进行处理器中要求特权态的操作,如I/O操作 `修改内存保护寄存器等.当用户程序需要进行这些操作时,必须通过系统调用进入管态,由操作系统代为完成.

45.计算机网络与分布式计算机系统的区别?

分布式系统定义是“存在一个能为用户自动管理资源的网络操作系统,由它调用完成用户任务所需要的资源,而整个网络像一个大的计算机系统一样对用户是透明的.”

分布式网络有以下5个特征:

- (1)系统中拥有多种通用的物理和逻辑资源,可以动态地给它们分配任务.
- (2)系统中分散的物理和逻辑资源通过计算机网络实现信息交换.
- (3)系统存在一个以全局方式管理系统资源的分布式操作系统.
- (4)系统中联网各计算机既合作又自治.
- (5)系统结构对用户是完全透明的.

计算机网络和分布式系统的共同点主要表现在:一般的分布式系统是建立在计算机网络之上的,因此分布式系统与计算机网络在物理结构上基本相同.

计算机网络与分布式系统的区别主要表现在:分布式操作系统与网络操作系统的设计思想是不同的,因此它们的结构 `工作方式与功能也是不同的.

分布式系统与计算机网络的主要区别不在它们的物理结构上,而是在高层软件上.分布式系

统是一个建立在网络之上的软件系统,这种软件保证了系统高度的一致性与透明性.分布式系统的用户不必关心网络环境中资源的分布情况,以及连网计算机的差异,用户的作业管理与文件管理过程是透明的.

计算机网络为分布式系统研究提供了技术基础,而分布式系统是计算机网络技术发展的高级阶段.

46.与文件系统相比,数据库文件有什么优点?

- (1)控制数据冗余.

数据库管理应尽可能地消除了冗余,但是没有完全消除,而是控制大量数据库固有的冗余.

- (2)保证数据的一致性

通过消除或者控制冗余,可降低不一致性产生的危险,如果数据项在数据库中只存储了一次,则任何对该值的更新均只需进行一次,而且新的值立即就被用户所获得.如果数据项不只存储了一次,而且系统意识到这点,系统将可以确保该项的所有拷贝都保持一致.

- (3)提高数据共享

数据库应该被有权限的用户共享.数据库管理系统的引入使更多的用户可以方便的共享更多的数据.新的应用程序可以依赖于数据库中已经存在的数据,而且只增加没有存储上的数据,而不用重新定义所有的数据需求.

47.说一下IP组播有哪些好处?

- (1)减少网络流量

组播可以减少网络流量,因为组播只需要一份数据包就可以同时发送给多个接收者,而不是每个接收者都发送一份数据包.

- (2)减少主机处理负担

组播可以减少主机的处理负担,因为主机只需要处理自己关心的数据包,而不是所有的数据包.

- (3)提高网络安全

组播可以提高网络安全,因为组播只发送给特定的接收者,而不是所有的接收者.

48.局部变量和全局变量能否重名?为什么?

能,局部会屏蔽全局.

49.计算机系统程序中进程和线程的差别?

线程是指进程内的一个执行单元,也可以是进程的调度实体.

与进程的区别:

- (1)地址空间:进程有自己独立的地址空间,而线程共享所属进程的地址空间.

- (2)资源拥有:进程是资源拥有的基本单位,线程不拥有系统资源,只拥有一点在运行中必不可少的资源.
- (3)独立性:进程是独立的基本执行单位,线程不是独立的基本执行单位,线程依赖于进程而存在.
- (4)系统开销:在创建或撤销进程时,由于系统都要为之分配或回收资源,导致系统的开销明显大于创建或撤销线程时的开销.

50.网络编程中设计并发服务器,使用的多线程和多进程的区别.

用多进程每个进程都有自己的地址空间,线程则共享地址空间.

- (1)速度:线程产生的速度快,线程间的通讯快、切换快,因为他们在一个地址空间里.
- (2)资源利用率:线程的资源利用率比较好也是因为他们在一个地址空间里.
- (3)同步问题:线程使用公共变量/内存时需要使用同步机制.

多线程的优势:

- (1)创建与环境切换开销小
- (2)同步的开销少
- (3)数据复制可通过“进程局部内存”

多线程的局限性:

- (1)性能损失
 - i. 单处理器、“计算任务繁重”的应用程序不会从多线程中获益
 - ii. 线程的创建和销毁会带来额外的开销
- (2)健壮性的降低
 - i. 线程之间接收到的“MMU保护”很少或没有
 - ii. 缺乏高精度的访问控制

51.点到点和端到端工作于哪一层,它的工作机制是什么?

点到点工作于网络层,端到端工作于传输层.

52.计算机中缓冲的定义,为什么引入它?

缓冲是一种临时存储区域,用于存放数据的一部分,以便等待处理器或设备的操作.引入缓冲的目的是为了解决生产者和消费者之间的速度不匹配问题,即生产者和消费者的速度不一致,为了使生产者和消费者之间能够协调工作,需要引入缓冲.

53.简述几种常见的设计模式

- (1)单例模式

单例模式是一种创建型设计模式,它保证一个类只有一个实例,并提供一个全局访问点.

- (2)工厂模式

工厂模式是一种创建型设计模式,它提供一个创建对象的接口,但是让子类决定实例化哪个类.

- (3)观察者模式

观察者模式是一种行为设计模式,它定义了一种一对多的依赖关系,让多个观察者对象同时监听某一个主题对象.

54.如何利用单纯形法求解线性规划最大值和最优解问题?

单纯形法是一种求解线性规划问题的方法,它通过不断地在可行解空间中移动,直到找到最优解.单纯形法的基本思想是从一个基本可行解出发,通过不断地移动,找到一个更优的基本可行解,直到找到最优解.

55.什么是知识表示,有哪些知识表示?

知识表示是研究用计算机表示知识的可行性、有效性的一般方法,是数据结构与系统控制结构的统一.知识表示的研究既要考虑知识的表示与存储,又要考虑知识的使用.知识表示可看成是一组事物描述的约定,是把人类知识表示成机器能处理的数据结构.

知识表示通常有直接表示,逻辑表示,产生式规则表示法,语义网络表示法,框架表示法,脚本表示法,过程表示,混合型知识表示方法.

56.什么是不确定性推理,不确定性推理的几种方法

推理的不确定性是由于知识不确定性的动态积累和传播过程所造成的.在推理的每一步中都需要综合各个证据、规则,引出结论,同时还要处理证据规则的不确定性.为此,整个过程要通过某种不确定性的度量,寻找尽可能符合客观世界的计算,最终得到结论的不确定性度量.

57.简述什么事SGD?

SGD是随机梯度下降算法,是梯度下降算法的一个拓展.机器学习反复出现的一个问题是好的泛化需要大的训练集,但大的训练集的计算代价也更大.机器学习算法中的代价函数通常可以分解成每个样本的代价函数的总和.随着训练集规模增长为数十亿的样本,进一步梯度也会消耗相当的时间.

随机梯度下降的核心是:梯度是期望.期望可使用小规模样本近似估计.具体而言,在算法的每一步,我们从训练集中均匀抽出一小批量样本 $B = \{x^{(1)}, \dots, x^{(m)}\}$.小批量的数目 m' 通常是一个相对较小的数,从一到几百.重要的是,当训练集大小 m 增长时, m' 通常是固定的.我们可能在拟合几十亿的样本时,每次更新计算只用到几百个样本

58.列举几种聚类的方法,K-means中的K由什么决定?

聚类方法有层次聚类 `K-means` `DBSCAN` `GMM`等.

K-means中的K由用户指定.

59.什么是缺页中断,影响缺页率的因素?

缺页(实际上是异常的一种,不属于中断)中断,是指要访问的页不在主存,需要操作系统将其调入主存后再进行访问

缺页率:在进行内存访问的时候,若所访问的页已在主存中,则称此次访问成功;若访问的页不在主存里,则称作缺页.

影响缺页率的因素:

- i. 分配给作业的主存块越多,则缺页率越低
- ii. 页面大,则缺页率低;页面小缺页率高.
- iii. 作业的访问局部性越强,缺页率越低.
- iv. 页面置换算法,算法的优劣决定了进程执行过程中缺页中断的次数.

60.什么是时间片,简述高响应比的调度算法的原理

时间片即CPU分配给各个程序的时间,每个线程被分配一个时间段,称作他的时间片,即该进程允许运行的时间,使各个程序从表面上看是同时进行的.如果在时间片结束时进程还在运行,则cpu将被剥夺并分配给另一个进程.

高响应比优先调度算法是一种对CPU中央处理器响应比的分配的一种算法,是介于先来先服务与短作业优先之间的这种方法,即考虑作业的等待时间又考虑运行时间.其中,响应比 = (等待时间+要求服务时间)/要求服务时间

61.简述进程调度的时机?

进程调度的时机有以下几种:

- (1) 进程创建或终止时
- (2) 进程阻塞或唤醒时
- (3) 进程时间片用完时
- (4) 进程优先级发生变化时
- (5) 进程由运行态转变为就绪态时
- (6) 进程由运行态转变为阻塞态时

62.简述分页与分段的区别

分页和分段是两种不同的内存管理方式.

- (1) 分页是将逻辑地址空间划分为固定大小的页,物理地址空间划分为相同大小的块,页和块的大小相同.分页是由硬件实现的,操作系统不需要管理页表,只需要管理页表基址寄存器和页表

长度寄存器.

- (2) 分段是将逻辑地址空间划分为若干个段,每个段的长度可以不同,物理地址空间划分为相同大小的块,段和块的大小不同.分段是由操作系统实现的,操作系统需要管理段表.

63.数据结构这门课中学习了什么算法?该算法的原理是什么?排序的过程是什么?

数据结构这门课中学习了冒泡排序 `选择排序 `插入排序 `希尔排序 `快速排序 `归并排序 `堆排序等算法.

冒泡排序的原理是:比较相邻的元素,如果第一个比第二个大,就交换他们两个.对每一对相邻元素作同样的工作,从开始第一对到结尾的最后一对.在这一点之后,最后的元素应该是最大的数.针对所有的元素重复以上的步骤,除了最后一个.时间复杂度为 $O(n^2)$,空间复杂度为 $O(1)$.

选择排序的原理是:首先在未排序序列中找到最小(大)元素,存放到排序序列的起始位置,然后,再从剩余未排序元素中继续寻找最小(大)元素,然后放到已排序序列的末尾.以此类推,直到所有元素均排序完毕.时间复杂度为 $O(n^2)$,空间复杂度为 $O(1)$.

插入排序的原理是:从第一个元素开始,该元素可以认为已经被排序.取出下一个元素,在已经排序的元素序列中从后向前扫描.如果该元素(已排序)大于新元素,将该元素移到下一位置.重复这个过程,直到找到已排序的元素小于或者等于新元素的位置.将新元素插入到该位置后.重复上述步骤.时间复杂度为 $O(n^2)$,空间复杂度为 $O(1)$.

64.分析Ping失败的原因?

1.目标主机不可达:当目标主机无法与当前主机建立网络连接时,Ping就会失败.可能的原因包括目标主机关目标主机所在网络故障或防火墙阻止了Ping请求.

2.网络故障:Ping失败有可能是由于网络设备 `链路或者路由器出现故障造成的.

3.ICMP协议限制:有些网络环境可能限制了ICMP协议的使用,这导致PING命令无法成功.