1. 在[软件开发](http://msd.misuland.com/pd/3255818066314926630)过程中，我们可以采用不同的过程模型，下列有关 增量模型描述正确的是（）

A. 是一种线性开发模型，具有不可回溯性 B. 把待开发的软件系统模块化，将每个模块作为一个增量组件，从而分批次地分析、设计、编码和测试这些增量组件 C. 适用于已有产品或产品原型（样品），只需客户化的工程项目 D. 软件开发过程每迭代一次，软件开发又前进一个层次

正确答案：B

这道题重复出现在《[京东](http://msd.misuland.com/pd/3255818100674658960)[2019](http://msd.misuland.com/pd/3255818100674661374)春招京东数据分析类试卷》第一题。

增量模型

增量模型也称为渐增模型，是把待开发的软件系统「模块化」，将每个模块作为一个增量组件，从而分批次地分析、设计、编码和测试这些增量组件。

优点： （1）将待开发的软件系统模块化，可以「分批次地提交软件产品」，各个阶段并不交付一个可运行的完整产品，而是交付满足客户需求的一个子集的可运行产品,使用户可以及时了解软件项目的进展。 （2）以组件为单位进行开发「降低了软件开发的风险」。一个开发周期内的错误不会影响到整个软件系统。 （3）「开发顺序灵活」。开发人员可以对组件的实现顺序进行优先级排序，先完成需求稳定的核心组件。当组件的优先级发生变化时，还能及时地对实现顺序进行调整。

缺点 （1）要求待开发的软件系统可以被模块化。如果待开发的软件系统很难被模块化，那么将会给增量开发带来很多麻烦。

来源：[百度](http://msd.misuland.com/pd/3255818100674661124)百科

「题目解析」 A. 具有可回溯性，可以返回修改； B. 正确； C. 适用于未开发的原型（样品）； D. 软件开发过程每迭代一次，软件开发又前进一个层次，属于螺旋模型的软件过程。

- 螺旋模型

强烈参考 《常用软件开发模型比较分析》（https://wenku.baidu.com/view/718cb107aaea998fcc220e70.html）

2. 下面有关值类型和引用类型描述正确的是（）？

A. 值类型的变量赋值只是进行数据复制，创建一个同值的新对象，而引用类型变量赋值，仅仅是把对象的引用的指针赋值给变量，使它们共用一个内存地址。 B. 值类型数据是在栈上分配内存空间，它的变量直接包含变量的实例，使用效率相对较高。而引用类型数据是分配在堆上，引用类型的变量通常包含一个指向实例的指针，变量通过指针来引用实例。 C. 引用类型一般都具有继承性，但是值类型一般都是封装的，因此值类型不能作为其他任何类型的基类。 D. 值类型变量的作用域主要是在栈上分配内存空间内，而引用类型变量作用域主要在分配的堆上。

正确答案：A B C

(求解析D)

 值类型

值类型就是现金，要用直接用；引用类型是存折，要用还得先去银行取现。

值类型 1.值类型变量都存储在栈中。2.访问值类型变量时，一般都是直接访问其实例。 3.每个值类型变量都有自己的数据副本，因此对一个值类型的变量的操作不会影响其他的变量。 4.复制值类型变量时，复制的是变量的值，而不是变量的地址。 5.值类型变量不能为null，必须具有一个确定的值。

引用类型 1.必须在托管堆中为引用类型变量分配内存。2.必须使用new关键字来创建引用类型变量。 3.在托管堆中分配的每个对象都有与之相关联的附加成员，这些成员必须被初始化。 4.引用类型变量是由垃圾回收机制来管理的。 5.多个引用类型变量都可以引用同一个对象，这种情形下，对一个变量的操作会影响另一个变量所引用的同一对象。 6.引用类型被赋值之前的值都是null。 来源：https://www.cnblogs.com/1ming/p/5227944.html

参考资料

https://www.nowcoder.com/test/question/done?tid=24816459&qid=313898#summary

https://www.cnblogs.com/1ming/p/5227944.html

3. 如何在多线程中避免发生死锁？

A. 允许进程同时访问某些资源。 B. 允许进程强行从占有者那里夺取某些资源。 C. 进程在运行前一次性地向系统申请它所需要的全部资源。 D. 把资源事先分类编号，按号分配，使进程在申请，占用资源时不会形成环路。

官方答案：A B C D 民间答案：B C D

死锁

在多道程序设计环境下，多个进程可能竞争一定数量的资源。一个进程申请资源，如果资源不可用，那么进程进入等待状态。如果所申请的资源被其他等待进程占有，那么该等待的进程有可能无法改变状态，这种情况下称之为死锁。

产生死锁的4个必要条件： 互斥条件：系统存在临界资源，存在一个资源每次只能被一个进程使用，若别的进程也要使用该资源，需要等待知道其占用者用完释放。 保持与等待条件：部分分配，允许进程在不释放其已经分得的资源的情况下请求并等待分配的资源 不可抢占条件：有些系统资源是不可抢占的，系即当某个进程已经获得这种资源后，系统是不能强行收回，其他进程也不能强行夺走，只能由自身使用完释放。 循环等待条件：若干个进程形成环形链，链中的每一个进程都在等待该链中下一个进程所占用的资源。 死锁的预防需要至少破坏死锁的4个必要条件之一，而死锁的避免不去刻意破坏4个必要条件，而是通过对资源的分配策略施加较少的限制条件，来避免死锁的产生。

「民间解析」

选项B，破坏不可抢占条件，属于死锁预防 选项C，破坏了保持与等待条件，属于死锁预防 选项D，破坏了循环等待条件，属于死锁预防 来源：https://www.nowcoder.com/questionTerminal/361381af0f2045d5874764e2c19a5b99?toCommentId=2528092 作者：尹沐君

参考资料

《如何在多线程中避免发生死锁》https://blog.csdn.net/weixin\_43113679/article/details/88365466

4.以下为求0到1000以内所有奇数和的[算法](http://msd.misuland.com/pd/3255818100674661178)，从中选出描述正确的算法（ ）

A. ①s=0；②i=1；③s=s+i；④i=i+2；⑤如果i≤1000，则返回③；⑥结束 B. ①s=0；②i=1；③i=i+2；④s=s+i；⑤如果i≤1000，则返回③；⑥结束 C. ①s=1；②i=1；③s=s+i；④i=i+2；⑤如果i≤1000，则返回③；⑥结束 D. ①s=1；②i=1；③i=i+2；④s=s+i；⑤如果i≤1000，则返回③；⑥结束

正确答案：A

5. 关于递归法的说法不正确的是（ ）

A. 程序结构更简洁 B. 占用CPU的处理时间更多 C. 要消耗大量的内存空间，程序执行慢，甚至无法执行 D. 递归法比递推法的执行效率更高

正确答案：D

「民间解析」

简单的说,递推是在借助前一个几经计算出来的结果去计算下一步的结果,以此来得到最终结果,有此可知递推并不需要保留太多现场信息,

而递归就不一样,虽然也是要借助前一步的结果,但这前一步结果往往刚开始是未知的,要一步一步递推下去,直到遇到终结条件,然后在一层一层的回归,直到回归到最上一层计算出结果,可见递归是包含两步的,一个递推下去,一个在回归

递归往往表达简单 ,但计算需要时空都比较大

参考资料

https://zhidao.baidu.com/question/295531510.html

https://www.nowcoder.com/questionTerminal/2c9a2595150044a0915df65e40704282?orderByHotValue=1&page=1&onlyReference=false

6. [字符串](http://msd.misuland.com/pd/3255818100674660788)”ABCD”和字符串”DCBA”进行比较，如果让比较的结果为真，应选用关系运算符（）

A. > B. < C. = D. >=

正确答案：B

字符串比较 优先判断对应字符的大小(ASCII码顺序)，A”为65；“a”为97；“0”为48；

「题目解析」 字符串”ABCD”和字符串”DCBA”，A小于D，则无需继续比较了，符号为 <。

7.下面是一段关于计算变量s的算法： ①变量s的初值是0 ②变量i从1起循环到n，此时变量s的值由下面的式子表达式计算 ③s=s+(-1)\*i ④输出变量s的值 这个计算s值的算法中，s的代数式表示是( )。

A. 1-2+3-4+„+(-1)n\*(n-1) B. 1-2+3-4+„+(-1)n-1\*n C. 1+2+3+4+…+(n-1)+n D. -1-2-3-4-…-n

正确答案：D

8.以下运算符中运算优先级最高的是（ ）

A. + B. OR C. > D. \

正确答案：D

运算符优先级 函数>算术>关系>逻辑

算数 指数（^）→ 取负（-）→ 乘浮点除（\*、/）→ 整除（\）→ 取模（Mod）→ 加、减（+、-）→ 连接（&）

关系 等于（＝）→不等于（！=）→小于（<）→大于（>）→ 小于等于（≤）→大于等于（≥）→Like→Is

逻辑 Not→And→Or→Xor→Eqv(等价)→Imp(蕴含)

9.采用哪种遍历方法可唯一确定一棵二叉树？（ ）

A. 给定一棵二叉树的先序和后序遍历序列 B. 给定一棵二叉树的后序和中序遍历序列 C. 给定先序、中序和后序遍历序列中的任意一个即可 D. 给定一棵二叉树的先序和中序遍历序列

正确答案：B D

二叉树的遍历方法 目前做了三套题，二叉树的遍历方法出现了三次。

| 套题 | 知识点 |
| --- | --- |
| 京东2019春招京东数据分析类试卷 | 第2题：根据前序和后续遍历，求中序遍历 |
| 网易2018校园招聘数据分析工程师笔试卷 | 第1题：已知中序遍历，求二叉树的所有可能 |
| 京东2019校招数据分析工程师笔试题 | 第9题：遍历方法确定唯一的二叉树 |

「题目解析」

前序和后序在本质上都是将父节点与子结点进行分离，但并没有指明左子树和右子树的能力，因此得到这两个序列只能明确父子关系，而不能确定一个二叉树。

给出中序遍历之后再给一个其他的遍历就能够确定了。

前序遍历+中序遍历

后序遍历+中序遍历 可以唯一确定一棵二叉树。

10.已知小顶堆：{51,32,73,23,42,62,99,14,24,3943,58,65,80,120}，请问62对应节点的左子节点是

A. 99 B. 73 C. 3943 D. 120

官方答案：B 个人答案：65

堆 来源：https://www.cnblogs.com/chengxiao/p/6129630.html

最小/大堆的构建方法，强烈参考 《最小堆 构建、插入、删除的过程图解》 https://blog.csdn.net/hrn1216/article/details/51465270

11. 若串S=”UP！UP！JD”，则其子串的数目

A. 33 B. 37 C. 39 D. 35

正确答案：B 个人答案：31

子串 串中任意个连续的字符组成的子序列称为该串的子串，是子集的概念，不是集合。空串也属于子串。

长度为n的字符串

串中字符均不相同 1、有n(n+1)/2 +1个子串； 2、非空子串：n（n+1）/2； 3、非空真子串：n（n+1）/2– 1；

串中字符出现重复 n(n+1)/2+1-重复个数

「题目解析」 UP！UP！JD，有8个字符，n = 8。 n(n+1)/2 + 1 = 37

但官方答案没有考虑重复的字符。若考虑重复字符， 1个字符子串重复个数 = 1 + 1 + 1 （U,P,!) 2个字符子串重复个数 = 1 + 1 (UP, P!) 3个字符子串重复个数 = 1 (UP!) 所以， 37 - 6 = 31

强烈参考

《求解字符串所包含子串的个数》 https://blog.csdn.net/zrcj0706/article/details/79409207

《求字符串的子串个数》 https://blog.csdn.net/xingchenhy/article/details/73555516

12. 一颗二叉树的叶子节点有5个，出度为1的结点有3个，该二叉树的结点总个数是？

A. 11 B. 12 C. 13 D. 14

正确答案：B

结点的度：结点拥有的子树的数目

叶子结点：度为0的结点

分支结点：度不为0的结点

树的度：树中结点的最大的度

 度和结点的关系

树中结点数 = 总[分叉](http://msd.misuland.com/pd/3255818066314921088)数 +1 即，度为0的节点数为度为2的节点数加1  n0 = n2 + 1

「题目解析」 已知，n0 = 5, n1 = 3

n = n0 + n1 + n2 n0 = n2 + 1 解得 n = 12。

13.以下哪种排序算法一趟结束后能够确定一个元素的最终位置？

A. 简单选择排序 B. 基数排序 C. 堆排序 D. 二路归并排序

官方答案：C 民间答案：A C

各种排序方法总结

简单选择排序，能够取出当前无序序列中最（小or大）值与第一位置的元素互换位置。

堆排序每趟总能选出一个最值位于根节点。

冒泡排序总是两两比较选出一个最值位于数组前面。

快速排序选出的枢轴在一趟排序中就位于了它最终的位置

插入排序（直接、二分）不一定会位于最终的位置，因为不确定后面插入的元素对于前面的元素是否产生影响。

希尔排序（本质也是插入排序）只在子序列中直接插入排序。所以不能确定。

二路归并排序除非在缓存区一次放入所有的序列（这样得不偿失），否则不能确定最终位置。

所以只有 简单选择排序、快速排序、冒泡排序、堆排序每一趟排序结束都能确定一个元素最终位置的方法有哪些

来源：https://zhidao.baidu.com/question/680635954191449972.html

14. 权值分别为9、3、2、8的结点，构造一棵哈夫曼树，该树的带权路径长度是？

A. 36 B. 40 C. 45 D. 46

正确答案：B

哈夫曼编码的知识点在《网易2018校园招聘数据分析工程师笔试卷》第二题出现过。 哈夫曼编码的理解和方法，强烈推荐参考《哈夫曼编码的理解(Huffman Coding)》（来源： https://blog.csdn.net/qq\_36653505/article/details/81701181 ）

「题目解析」 带权路径长度 = (2+3)x3 + 8 x 2 + 9 x 1 = 40

15. 在（）中，只要指出表中任何一个结点的位置，就可以从它出发依次访问到表中其他所有结点。

A. 线性单链表 B. 双向链表 C. 线性链表 D. 循环链表

正确答案：D

单向列表 图片来源：https://www.cnblogs.com/skywang12345/p/3561803.html

双向列表 图片来源：https://www.cnblogs.com/skywang12345/p/3561803.html

循环列表 图片来源：https://zhidao.baidu.com/question/1732464800536322507.html

「题目解析」

单向链表： 一个节点包含2个信息：值，下一个节点，所以，对于任何一个节点，只能知道它的下一个节点，末尾节点的下一个节点不存在；

双向链表：每个结点设置有两个指针，一个指向其前驱，一个指向其后继，这样从任意一个结点开始，既可以向前查找，也可以向后查找，在结点的访问过程中一般从当前结点向链尾方向扫描，如果没有找到，则从链尾向头结点方向扫描，这样部分结点就要被遍历两次，因此不符合题意。

循环链表的最后一个结点的指针域指向表头结点，所有结点的指针构成了一个环状链，只要指出表中任何一个结点的位置就可以从它出发访问到表中其他所有的结点。

强烈参考

http://www.100xuexi.com/question/questiondetail.aspx?tid=2987&pid=8422155e-564a-4474-9cd6-4be42e95ba8d&qid=7639e754-531e-4bea-904e-226ffca2e0cb

《数组、单链表和双链表介绍 以及 双向链表的C/C++/Java实现》https://www.cnblogs.com/skywang12345/p/3561803.html

https://zhidao.baidu.com/question/1732464800536322507.html

16. 网络管理员把优盘上的[源代码](http://msd.misuland.com/pd/3255818100674660594)给[程序员](http://msd.misuland.com/pd/3255818100674660240)参考，但要防止程序误删除或修改，以下正确的加载方式是（ ）

A. mount -o defaults /dev/sdb1 /tools B. mount -r /dev/sdb1 /tools C. mount -o ro /dev/sdb1 /tools D. mount -o ro /dev/sdb /tools

正确答案：B C

linux磁盘管理

任何块设备都不能直接访问，需挂载在目录上访问 「挂载」: 将额外文件系统与根文件系统某现存的目录建立起关联关系，进而使得此目录做为其它文件访问入口的行为（挂载的设备必须有文件系统） mount [-选项] DEVICE（设备） MOUNT\_POINT（挂载点） （1）device ：指明要挂载的设备 ① 设备文件：例如/dev/sda5 （2）挂载选项- -r：readonly ，只读挂载，不能进行其他操作，和权限无关，介质只能读，多用于冷备份。在mount 查询时时显示 ro -o options：( 挂载文件系统的选项) ，多个选项使用逗号分隔 -o ro：只读 rw：读写(defaults) 来源：https://www.linuxidc.com/Linux/2017-11/148277p3.htm

「民间解析」

A- 默认选项–错误 D- sdb 和sdb1的差别。sdb表示整个SD卡设备名 sdb1表示SD卡的第一个分区。相应的，如果还有一个分区的话，表示为sdb2 用fdisk -l命令查看U盘的盘符。 一般为/dev/sdb1 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313912#summary 作者：weixy18

17. 下列有关软连接描述正确的是

A. 与普通文件没什么不同，inode 都指向同一个文件在硬盘中的[区块](http://msd.misuland.com/pd/3255818100674661228) B. 不能对目录创建软链接 C. 保存了其代表的文件的绝对路径，是另外一种文件，在硬盘上有独立的区块，访问时替换自身路径 D. 不可以对不存在的文件创建软链接

官方答案：B 民间答案：C

软连接

硬连接 若一个 inode 对应多个文件名，则称这些文件为硬链接。换言之，硬链接就是同一个文件使用了多个别名。

软连接 软链接与硬链接不同，若文件用户数据块中存放的内容是另一文件的路径名的指向，则该文件就是软连接。软链接就是一个普通文件，只是数据块内容有点特殊。软链接有着自己的 inode 号以及用户数据块。因此软链接的创建与使用没有类似硬链接的诸多限制。 来源：https://www.cnblogs.com/Rosanna/p/3598929.html

「民间解析」

A错误，后半句说的是硬链接，硬链接是共同拥有同一个inode。 B错误，软链接可对文件或目录创建 D错误，可对不存在的文件或目录创建软链接； 来源：https://www.nowcoder.com/questionTerminal/bafbec5b3fbb4dda8644b10c955f620e

参考资料 《硬链接、软链接和inode 》 https://www.cnblogs.com/Rosanna/p/3598929.html

18. Ext3日志文件系统的特点是：

A. 高可用性 B. 数据的完整性 C. 数据转换快 D. 多日志模式

正确答案：A B C D

EXT3

EXT3是第三代扩展文件系统（英语：Third extended filesystem，缩写为ext3），是一个日志文件系统，常用于Linux[操作系统](http://msd.misuland.com/pd/3255818100674658804)。

Ext3日志文件系统特点 1、高可用性 系统使用了ext3文件系统后，即使在非正常关机后，系统也不需要检查文件系统。宕机发生后，恢复ext3文件系统的时间只要数十秒钟。 2、数据的完整性 ext3文件系统能够极大地提高文件系统的完整性，避免了意外宕机对文件系统的破坏。 3、文件系统的速度 尽管使用ext3文件系统时，有时在存储数据时可能要多次写数据，但是，从总体上看来，ext3比ext2的性能还要好一些。这是因为ext3的日志功能对磁盘的驱动器读写头进行了优化。 4、数据转换 由ext2文件系统转换成ext3文件系统非常容易，只要简单地键入两条命令即可完成整个转换过程，用户不用花时间备份、恢复、格式化分区等。 5、多种日志模式 Ext3有多种日志模式，一种工作模式是对所有的文件数据及metadata（定义文件系统中数据的数据,即数据的数据）进行日志记录（data=journal模式）；另一种工作模式则是只对metadata记录日志，而不对数据进行日志记录，也即所谓data=ordered或者data=writeback模式。 来源：https://baike.baidu.com/item/Ext3/822591?fr=aladdin#4

19. DHCP是动态主机配置协议的简称,其作用是

A. 动态分配磁盘资源 B. 动态分配内存资源 C. 为网络中的主机分配IP地址 D. 为集群中的主机分配IP地址

正确答案：C

DHCP

动态主机设置协议（英语：Dynamic Host Configuration Protocol，DHCP）是一个局域网的网络协议，使用UDP协议工作，主要有两个用途： （1）用于内部网或网络服务供应商自动分配IP地址； （2）给用户用于内部网管理员作为对所有[计算机](http://msd.misuland.com/pd/3255818100674660634)作中央管理的手段。 来源：https://baike.baidu.com/item/DHCP/218195?fr=aladdin

20. 以下命令可以用于获取本地ip地址的是：

A. ifconfig B. uptime C. top D. netstat

正确答案：A

「题目解析」

A. ifconfig：获取本地ip地址； B.uptime：显示系统已经运行了多长时间，它依次显示下列信息：当前时间、系统已经运行了多长时间、目前有多少登陆用户、系统在过去的1分钟、5分钟和15分钟内的平均负载； C. top：查看linux系统的CPU、内存、运行时间、交换分区、执行的线程等信息； D.netstat：列出系统上所有的网络套接字连接情况，包括 tcp, udp 以及 unix套接字，另外它还能列出处于监听状态（即等待接入请求）的套接字。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313916#summary

21. 以下命令可以用于获取本机cpu使用率的是：

A. ifconfig B. uptime C. top D. netstat

正确答案：C 参考第20题解析。

22. 以下命令用于设置环境变量的是：

A. export B. cat C. echo D. env

正确答案：A

参考《京东2019春招京东数据分析类试卷》第4题解析。

关于环境变量的命令 export： 设置环境变量 echo：查看是否成功 env：显示所有的环境变量 set：显示所有本地定义的Shell变量 unset：清除环境变量

23. 下列对TCP/IP结构及协议分层不正确的是：

A. 网络接口层：Wi-Fi、ATM 、GPRS、EVDO、HSPA。 B. 网际层：IP、ICMP、IGMP 。 C. 传输层：TCP、UDP、TLS、ssh。 D. FTP、TELNET、DNS、SMTP.

正确答案：C

TCP/IP协议

TCP/IP协议（传输控制协议/[互联网](http://msd.misuland.com/pd/3255818100674660490)协议）不是简单的一个协议，而是一组特别的协议，包括：TCP，IP，UDP，ARP等，这些被称为子协议。在这些协议中，最重要、最著名的就是TCP和IP。因此，大部分网络管理员称整个协议族为“TCP/IP”。

来源：https://www.jianshu.com/p/9f3e879a4c9c

参考资料

https://www.jianshu.com/p/9f3e879a4c9c

https://www.cnblogs.com/dadadechengzi/p/7999371.html

24. 以下哪种设备工作在数据链路层？

A. 中继器 B. 集线器 C. 交换机 D. 路由器

正确答案：C

OSI七层模型涉及的主要设备

物理层的主要设备：中继器、集线器。

数据链路层主要设备：二层交换机、网桥

网络层主要设备：路由器

25. 打电话使用的数据传输方式是（），手机上网使用的数据传输方式是（）？

A. 电路交换，电路交换 B. 电路交换，分组交换 C. 分组交换，分组交换 D. 分组交换，电路交换

正确答案：B

「电路交换」是以电路为目的的交换方式，即通信双方要通过电路建立联系，建立后没挂断则电路一直保持，实时性高。

「分组交换」是把信息分为若干分组，每个分组有分组头含有选路和控制信息，可以到达收信方，但是不能即时通信 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313921#summary

26. 后退N帧协议的发送窗口大小是（），接收窗口的大小是（）？

A. =1,=1 B. =1,>1 C. >1,>1 D. >1,=1

正确答案：D

后退N帧ARQ协议对传统的自动重传请求（ARQ,Automatic Repeat reQues）进行了改进，从而实现了在接收到ACK之前能够连续发送多个数据包。 注：ACK：在TCP/IP协议中，如果接收方成功的接收到数据，那么会回复一个ACK数据。通常ACK信号有自己固定的格式,长度大小,由接收方回复给发送方。

接收窗口接收到分组就向前移动。发送窗口接受到接收窗口发送的ACK才会向前移动。

后退N帧协议的最大发送窗口为2n-1(其中n为帧号的位数)，最小为1， 接收窗口大小始终为1。

27.TCP释放连接第二次挥手时ACK（）,第三次挥手时ACK（）?

A. 不存在，不存在， B. 不存在，值是1 C. 值是1，值是1 D. 值是1，不存在

答案：C

「民间解析」

TCP连接建立阶段：

第一次握手：客户端的应用进程主动打开，并向客户端发出请求报文段。其首部中：SYN=1,seq=x。

第二次握手：[服务器](http://msd.misuland.com/pd/3255818100674660546)应用进程被动打开。若同意客户端的请求，则发回确认报文，其首部中：SYN=1,ACK=1,ack=x+1,seq=y。

第三次握手：客户端收到确认报文之后，通知上层应用进程连接已建立，并向服务器发出确认报文，其首部：ACK=1,ack=y+1。当服务器收到客户端的确认报文之后，也通知其上层应用进程连接已建立。

连接释放阶段：

第一次挥手：数据传输结束以后，客户端的应用进程发出连接释放报文段，并停止发送数据，其首部：FIN=1,seq=u。

第二次挥手：服务器端收到连接释放报文段之后，发出确认报文，其首部：ack=u+1,seq=v。此时本次连接就进入了半关闭状态，客户端不再向服务器发送数据。而服务器端仍会继续发送。

第三次挥手：若服务器已经没有要向客户端发送的数据，其应用进程就通知服务器释放TCP连接。这个阶段服务器所发出的最后一个报文的首部应为：FIN=1,ACK=1,seq=w,ack=u+1。

第四次挥手：客户端收到连接释放报文段之后，必须发出确认：ACK=1,seq=u+1,ack=w+1。 再经过2MSL(最长报文端寿命)后，本次TCP连接真正结束，通信双方完成了他们的告别。因此应选B 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313923#summary 作者：喵星人拯救地球

28. TCP协议的拥塞控制就是防止过多的数据注入到网络中，这样可以使网络中的路由器或链路不致过载。常用的方法有:

A. 慢启动、窗口滑动 B. 慢开始、拥塞控制 C. 快重传、快恢复 D. 快开始、快恢复

正确答案：B C

TCP拥塞控制 慢启动、拥塞避免、快重传、快启动

强烈参考 《TCP拥塞控制-慢启动、拥塞避免、快重传、快启动》https://blog.csdn.net/jtracydy/article/details/52366461

29. 对于京东商城高流量访问，预防Ddos的方法可以有？

A. 限制同时打开SYN半链接的数目。 B. 缩短SYN半链接的Time out 时间。 C. 关闭不必要的服务。 D. 限制客户端请求服务器时长。

正确答案：A B C

Ddos

分布式拒绝服务攻击(Distributed Denial of Service，简称DDoS)是指处于不同位置的多个攻击者同时向一个或数个目标发动攻击，或者一个攻击者控制了位于不同位置的多台机器并利用这些机器对受害者同时实施攻击。由于攻击的发出点是分布在不同地方的，这类攻击称为分布式拒绝服务攻击，其中的攻击者可以有多个。 来源：https://baike.baidu.com/item/分布式拒绝服务攻击/3802159?fromtitle=DDOS&fromid=444572&fr=aladdin

SYN

同步序列编号（Synchronize Sequence Numbers）。是TCP/IP建立连接时使用的握手信号。在客户机和服务器之间建立正常的TCP网络连接时，客户机首先发出一个SYN消息，服务器使用SYN+ACK应答表示接收到了这个消息，最后客户机再以ACK消息响应。这样在客户机和服务器之间才能建立起可靠的TCP连接，数据才可以在客户机和服务器之间传递。 在[黑客](http://msd.misuland.com/pd/602985665654689792)攻击事件中，SYN攻击是最常见又最容易被利用的一种攻击手法。 来源：https://baike.baidu.com/item/SYN

30.重复的数据，会增加磁盘空间的占有率，延长操作数据的时间。可以使用规范化处理数据冗余，以下对符合第一范式的表述正确的是：

A. 非键属性和键（主键）属性间没有传递依赖 B. 非键属性和键（主键）属性间没有部分依赖 C. 表中不应该有重复组。列重复拆成另外一张表；行重复拆成多行 D. 一个表中的列值与其他表中的主键匹配

正确答案：C

范式

「第一范式（1NF）」 指[数据库](http://msd.misuland.com/pd/3255818100674661272)表的每一列都是不可分割的基本数据项，同一列中不能有多个值，即实体中的某个属性不能有多个值或者不能有重复的属性。

「第二范式（2NF）」 第二范式（2NF）是在第一范式（1NF）的基础上建立起来的，第二范式就是属性完全依赖于主键。

「第三范式（3NF）」 满足第三范式（3NF）必须先满足第二范式（2NF）。简而言之，第三范式（3NF）要求一个数据库表中不包含已在其它表中已包含的非主关键字信息。或者说，第三范式就是属性不依赖于其它非主属性。 来源：https://www.cnblogs.com/ktao/p/7775100.html

强烈参考 《第一、第二、第三范式之间的理解和比较》可参考其中的实例讲解；https://www.cnblogs.com/ktao/p/7775100.html

31. 以下哪条SQL语句可以返回table1中的全部的key：

A. select tabel1.key from table1 join tabel2 on table1.key=table2.key B. select tabel1.key from table1 right outer join tabel2 on table1.key=table2.key C. select tabel1.key from table1 left semi join tabel2 on table1.key=table2.key D. select tabel1.key from table1 left outer join tabel2 on table1.key=table2.key

正确答案：D

 来源网络。

举一反三

关于左链接的SQL实例，可参考《牛客网SQL实战二刷 | Day1》第五题。

32.以下关于Mysql数据库引擎ＭyISAM的描述错误的是？

A. 支持行锁 B. 如果表主要是用于插入新记录和读出记录，那么选择MyISAM引擎能实现处理高效率 C. 在执行查询语句（SELECT）前，会自动给涉及的所有表加读锁，在执行更新操作（UPDATE、DELETE、INSERT等）前，会自动给涉及的表加写锁 D. 不支持事务

正确答案：A

「民间解析」

「MyISAM引擎」 是一种非事务性(不支持事务)的引擎，提供高速存储和检索，以及全文搜索能力，适合数据仓库等查询频繁的应用。

MyISAM和Innodb的区别： Innodb支持事务处理与外键和行级锁，而MyISAM不支持。所以MyISAM往往就 容易被人认为只适合在小项目中使用。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313928#summary 作者：LGXZ

33.Mysql中表user的建表语句如下，

CREATE TABLE `user` (

`id` bigint(20) NOT NULL AUTO\_INCREMENT COMMENT '主键Id',

`name` varchar(255) DEFAULT NULL COMMENT '名称',

`age` int(11) DEFAULT NULL COMMENT '年龄',

`address` varchar(255) DEFAULT NULL COMMENT '地址',

`created\_time` datetime DEFAULT NULL COMMENT '创建时间',

`updated\_time` datetime DEFAULT NULL COMMENT '更新时间',

PRIMARY KEY (`id`),

KEY `idx\_com1` (`name`,`age`,`address`)

) ENGINE=InnoDB DEFAULT CHARSET=utf8 COMMENT='用户表';

以下哪个查询语句没有使用到索引idx\_com1？

A. select \* from user where name=‘张三’ and age = 25 and address=‘北京大兴区’; B. select \* from user where name=‘张三’ and address=‘北京大兴区’; C. select \* from user where age = 25 and address=‘北京大兴区’; D. select \* from user where address=‘北京大兴区’ and age = 25 and name=‘张三’

正确答案：C

「民间解析」

索引在使用的时候要遵守「最左原则」，这里的复合索引最左字段为name，在创建 idx\_com1(name,age,address)索引的时候，实际上是创建了(name)，（name，age），（name,age,address）三种索引，C选项里面没有关联到name字段，因此选择C。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313929#summary 作者：Jessidea

34. DELETE和TRUNCATE TABLE都是删除表中的数据的语句，它们的不同之处描述正确的是：

A. TRUNCATE TABLE比DELETE的速度快 B. 在删除时如果遇到任何一行违反约束（主要是外键约束），TRUNCATE TABLE仍然删除，只是表的结构及其列、约束、索引等保持不变，但DELETE是直接返回错误 C. 对于被外键约束的表，不能使用TRUNCATE TABLE，而应该使用不带WHERE语句的DELETE语句。 D. 如果想保留标识计数值，要用DELETE，因为TRUNCATE TABLE会对新行标志符列使用的计数值重置为该列的种子

正确答案：A B C D

强烈参考

https://www.cnblogs.com/Herzog3/p/5955339.html

https://blog.csdn.net/zhichunqi/article/details/80780942

35. 用户表中有两列name/country。现在要查询用户表中每个国家(country)的用户人数，应使用以下哪个语句

A, select count(\*) from users group by country B. select name from users where country = ‘xx’ C. select count(country) from users D. select country from users

正确答案：A

「题目解析」 用GROUP BY语句分组，COUNT（）函数计数。 实例可参考《牛客网SQL实战二刷 | Day2》第7题。

36. 下列程序打印结果为（ ）

import re

m = re.search('[0-9]','a1b2c3d4')print(m.group(0))

A. 1 B. 1234 C. None D. 其他几项都不对

正确答案：A

正则表达式

「题目解析」

re.search('[0-9]','a1b2c3d4')

意思为在字符串'a1b2c3d4'中寻找数字0-9。

m.group(0) 表示输出匹配到的结果。

故第一个匹配到的数字是1。

37. 下列程序打印结果为（ ）

import datetime

t1 = datetime.datetime(2017,10,10,21,40)

t2 = datetime.datetime(2017,10,8,23,40)

tt1 = datetime.timedelta(seconds = 1200)

tt2 = datetime.timedelta(weeks = 3)

print(t1 - t2)

A. 600 B. 3600 C. 1 day, 02:00:00 D. 1 day, 22:00:00

正确答案：D

38. 将Person表中Name字段为"Lilei"的AGE字段递增1，可以使用哪个SQL语句（ ）

A. UPDATE Person SET AGE = AGE + 1 WHERE Name = 'Lilei'

B. Alter Person SET AGE = AGE + 1 WHERE Name = 'Lilei'

C. UPDATE Person SET AGE = AGE + 1 WHERE Name = Lilei

D. SET AGE = AGE + 1 FROM Person WHERE Name = 'Lilei'

正确答案：A

39. 将编码为gbk的字符串s转码为utf-8编码，以下操作正确的是

A. s.encode('utf-8')

B. s.decode('gbk')

C. s.encode('gbk').decode('utf-8')

D. s.decode('gbk').encode('utf-8')

正确答案：D decode解码，encode编码。

40. 以下关于range/xrange说法错误的是：

A. range返回的是list B. xrange返回的是list C. 生成大的数字序列时，适宜用range D. 生成大的数字序列时，适宜用xrange

正确答案：B C

range

「range」 range([start,] stop[, step])，根据start与stop指定的范围以及step设定的步长，生成一个序列。

「xrange」 用法与range完全相同，所不同的是生成的不是一个数组，而是一个生成器。

比较 要生成很大的数字序列的时候，用xrange会比range性能优很多，因为不需要一上来就开辟一块很大的内存空间，这两个基本上都是在循环的时候用。

强烈参考 《Python中range和xrange的区别》 https://www.cnblogs.com/xiezhiyang/p/6613094.html

41.python函数中支持\*args和\*\*kwargs。关于二者的区别以下说法错误的是：

A. args要位于kwargs之前 B. kwargs是将相应的传值以字典形式呈现 C. args是将相应的传值以list的形式呈现 D. kwargs表示关键字参数

官方答案：C 民间答案：C D

args kwargs

普通参数，即在调用函数时必须按照准确的顺序来进行参数传递。

默认参数，即参数含有默认值，在调用函数时可以进行参数传递，若没有进行参数传递则使用默认值，要注意，默认参数必须在普通参数的右侧（否则解释器无法解析）。

元组参数，即 \*args，参数格式化存储在一个元组中，长度没有限制，必须位于普通参数和默认参数之后。

字典参数，即 \*\*kwargs，参数格式化存储在一个字典中，必须位于参数列表的最后面。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313937#summary

注意点：参数arg、\*args、\*\*kwargs三个参数的位置必须是一定的。必须是(arg,\*args,\*\*kwargs)这个顺序，否则程序会报错。 来源：https://www.cnblogs.com/yunguoxiaoqiao/p/7626992.html

强烈参考 《Python中\*args和\*\*kwargs的区别》 https://www.cnblogs.com/yunguoxiaoqiao/p/7626992.html

42. 系统管理员编写扫描临时文件的shell程序tmpsc.sh, 测试该程序时提示拒绝执行，解决的方法有（ ）

A. chmod 644 tmpsc.sh B. chmod 755 tmpsc.sh C. chmod a+x tmpsc.sh D. chmod u+x tmpsc.sh

正确答案：B C D

「题目解析」

chmod 用3个数字来表达对 用户（文件或目录的所有者），用户组（同组用户），其他用户 的权限： 如：chmod 777 /test 数字7是表达同时具有读，写，执行权限：

读取–用数字4表示；

写入–用数字2表示；

执行–用数字1表示；

按照规则，如你想设置/test目录的权限为： 对用户可读可写：4（读取）+ 2（写入）= 6 ； 对用户组可读可执行：4（读取）+ 1（执行）= 5 ； 对其他用户仅可读：4（读取）； 这样就可以用命令： chmod 654 /test

chmod是权限管理命令change the permissions mode of a file的缩写；u代表所有者user；x代表执行权限；+ 表示增加权限。

chmod u+x file.sh就表示对当前目录下的file.sh文件的所有者增加可执行权限。

a+x 是给所有人加上可执行权限，包括所有者，所属组，和其他人

o+x 只是给其他人加上可执行权限

来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313938#summary

43. 在bash编程中,算术比较大于、大于等于的运算符是（ ）

A. > B. >= C. ge D. gt

正确答案：C D ge -> greater or equal gt -> greater than

44. echo expr 3/4的执行结果

A. 0 B. 1 C. 0.75 D. 3/4

正确答案：D (求解析～)

计算整数表达式的运算结果 格式：expr 变量1 运算符 变量2 …[运算符 变量n] • 「expr的常用运算符」 加法运算：+ 减法运算： - 乘法运算： \* 除法运算： / 求模（取余）运算： %

45. 文件目录data当前权限为rwx — ---，只需要增加用户组可读权限，但不允许写操作，具体方法为：

A. chmod +050 data B. chmod +040 data C. chmod +005 data D. chmod +004 data

正确答案：A

参考第42题

chmod 用3个数字来表达对 用户（文件或目录的所有者），用户组（同组用户），其他用户 的权限： 如：chmod 777 /test 数字7是表达同时具有读，写，执行权限：

读取–用数字4表示；

写入–用数字2表示；

执行–用数字1表示；

46. bash脚本文件一般第一行开头是

A. // B. ## C. #! D. #/

正确答案：C

47. 如何获取上一条命令执行的返回码

A. $! B. $0 C. $? D. $#

正确答案：C

shell 命令

$$ Shell本身的PID（ProcessID）

$! Shell最后运行的后台进程的PID

$? 最后运行的命令的结束代码

$- 使用Set命令设定的Flag一览

$\* 所有参数列表。以$1 $2 … $n的形式输出所有参数。

@所有参数列表。没有被引用时，与 @所有参数列表。没有被引用时，与没有区别。被引用时，"$“相当于”$1 $2 … n&quot;， n"，＠相当于"$1",“ 2&quot; …&quot; 2"…"n”

$# 添加到Shell的参数个数

$0 Shell本身的文件名 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313943#summary 作者：喵星人拯救地球

48. Shell 脚本（shell script），是一种为 shell 编写的脚本程序。现有一个test.sh文件，且有可执行权限，文件中内容为：

#!/bin/bash

aa='Hello World !'

请问下面选项中哪个能正常显示Hello World !

A. sh test.sh >/dev/null 1 && echo $aa B. ./test.sh >/dev/null 1 && echo $aa C. bash test.sh >/dev/null 1 && echo $aa D. . ./test.sh >/dev/null 1 && echo $aa

正确答案：D

「民间解析」

ABC都是开子shell，不影响当前环境中的变量 . ./test.sh等于source ./test.sh直接影响当前环境下的变量 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313944#summary 作者：Dominic.z

（求解析～）

49.以下哪个命令是将标准输出和错误重定向到a.txt文件

A. &>a.txt B. &|a.txt C. a.txt < & D. a.txt | &

正确答案：A

标准输出重定

如果希望把标准输出重定向到文件中，可以用>filename。在下面的例子中，ls命令的所有输出都被重定向到ls.out文件中： $ ls >ls.out 来源：https://www.cnblogs.com/softidea/p/3965093.html

参考资料

https://www.cnblogs.com/softidea/p/3965093.html

https://www.cnblogs.com/selectztl/p/9477988.html

50. 假设一种基因同时导致两件事情，一是使人喜欢抽烟，二是使这个人和肺癌就是（）关系，而吸烟和肺癌则是（）关系。

A. 因果；相关 B. 相关；因果 C. 并列；相关 D. 因果；并列

正确答案：A

51. 若一个学习器的ROC曲线被另外一个学习器低的曲线完全“包住”，则断言后者的性能优于前者；若两个学习器的曲线出现交叉，该如何处理最为合适？

A. 比较ROC曲线线上的面积 B. 使用AUC进行比较 C. 目测进行判断 D. 通过其他方法判断两个学习器的优劣

正确答案：B

AUC是ROC曲线围住的面积，越大越好。

52. 关于随机森林的训练过程下列描述正确的是：

A. 样本扰动 B. 属性扰动 C. 样本扰动并且属性扰动 D. 不存在扰动现象

正确答案：C

「题目解析」

样本扰动： 例如bagging算法的输入样本重采样（有放回的采样出与原始样本容量相同的新样本）对输入样本扰动敏感的不稳定学习器例如决策树、神经网路等适合此种方式，但是要注意，有的基学习器对该扰动不敏感，例如，线性回归、支持向量机、朴素贝叶斯、K近邻该类学习器称为稳定学习器。

属性扰动： 最经典的例子是随机森林，传统的决策树是从某个特征所有的属性值中选出最优的切分点进行数据样本划分，但是随机森林是，随机选取某个特征部分属性值作为属性子集，并从该子集中选出最优的切分点。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313949#summary 作者：喵星人拯救地球

53. ( )的目的缩小数据的取值范围,使其更适合于数据挖掘算法的需要,并且能够得到 和原始数据相同的分析结果。

A. 数据清洗 B. 数据集成 C. 数据变换 D. 数据归约

正确答案：D

数据集成：将多个数据源中的数据合并并存放在一个一致的数据仓库中。 数据变换：将数据转换或统一成适合挖掘的形式。 数据清洗：将重复、多余的数据筛选清除，将缺失的数据补充完整，将错误的数据纠正或者删除，最后整理成为我们可以进一步加工、使用的数据。 数据归约：在尽可能保持数据原貌的前提下，最大限度地精简数据量。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313950#summary 作者：喵星人拯救地球

参考资料

数据清洗：https://www.cnblogs.com/charlotte77/p/5606926.html

数据变换：https://blog.csdn.net/sysstc/article/details/84532396

数据规约：https://baike.baidu.com/item/数据归约/1639875?fr=aladdin

54. 假设属性income的最大最小值分别是12000元和98000元。利用最大最小规范化的方法将属性的值映射到0至1的范围内。对属性income的73600元将被转化为（）

A. 0.821 B. 1.224 C. 1.458 D. 0.716

正确答案：D

「题目解析」 (73600-12000)/(98000-12000) = 0.716

55. 下列关于脏数据的说法中，正确的是()

A. 格式不规范 B. 编码不统一 C. 意义不明确 D. 与实际业务关系不大 E. 数据不完整

正确答案：A B C D E

56. 图像挖掘中常用卷积神经网络（DNN）作为基础结构，以下关于卷积操作 （conv）和池化 （ pooling）的说法正确的是？

A. conv基于平移不变性，pooling基于局部相关性 B. conv和pooling都基于平移不变性 C. conv基于局部相关性，pooling基于平移不变性 D. conv和pooling都基于局部相关性

正确答案：A

平移不变性： 对于同一张图及其平移后的版本，都能输出同样的结果

局部相关性： 池化层利用局部相关性，对图像进行下采样，可以减少数据处理量同时保留有用信息，相当于图像压缩。 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313953#summary 作者：喵星人拯救地球

57. 京东展开足球比赛，已知：joy队和Jder队。假设65%的比赛joy队胜出,剩余的比赛Jder队获胜。joy队获胜的比赛中只有30%是在Jder队的主场,而Jder队取胜的比赛中75%是主场获胜。如果下一场比赛在Jder队的主场进行Jder队获胜的概率为

A. 0.75 B. 0.35 C. 0.4678 D. 0.5738

官方答案：C 民间答案：D

贝叶斯公式

58.为了培养员工和子女的亲密度，京东开展亲子活动。活动中进行分组，一组有4对亲子（父子，母女，父女，母子），分组后同一组坐在同一张圆桌旁。活动中规定，孩子旁边只能是其他小孩或者自己父母 ，那么4对亲子在圆桌上有几种坐法？

A. 144 B. 240 C. 576 D. 480

答案：D

「民间解析」

不能让孩子左右两边都是父母，否则旁边至少有一个不是自己父母。只能四个父母相邻，或两对父母相邻。

若四个父母相邻，两孩子固定，两孩子随意：A(4,4)\*A(2,2)\*8=384

若两个父母相邻，两对父母必须面对面，孩子均在自己父母旁边。相当于，只安排四个孩子的座位A（4，4），再旋转一圈\*8，由于堆成，再除以2：A(4,4)\*8/2=96 参考：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313955#summary 作者：喵星人拯救地球

59.有三道门d1，d2，d3，兔子出现在各道门后的概率相同，选手猜中兔子在哪道门后面，就能赢得大奖。假设选手现在选门d1，主持人这时候去观察另外两扇门d2,d3，并明确告诉选手d3后面没有兔子。此时选手可以更改自己的选择，以便最大可能赢得奖励。以下说法正确的是？

A. 选手应该选择d1，且猜中概率为1/2 B. 选手应该选择d1，且猜中概率为2/3 C. 选手应该选择d2，且猜中概率为1/2 D. 选手应该选择d2，且猜中概率为2/3

正确答案：D

「民间解析」

答案的前提是，该游戏主持人一定会打开一扇空门。如果选手所选门后没有兔子，主持人只会打开剩下两扇门中，没兔子的门；如果选手所选门后有兔子，主持人随机打开剩下两扇门中的一扇门。 d1,d2,d3门后有兔子概率为P(d1)=P(d2)=P(d3)=1/3。现在已知选手选择d1，主持人打开d3为空，

若兔子在d1，则主持人打开d3的概率P(开d3｜d1)=1/2

若兔子在d2，则主持人打开d3的概率P(开d3｜d2)=1

若兔子在d3，则主持人打开d3的概率率P(开d3｜d3)=0

因此，主持人打开d3的概率为P(开d3)=1/6+1/3+0=1/2

在主持人打开d3的条件下，d1后有兔子概率P(d1｜开d3)=P(d1)P(开d3｜d1)/P(开d3)=1/31/2\*2=1/3

在主持人打开d3的条件下，d2后有兔子概率P(d2｜开d3)=P(d2)P(开d3｜d2)/P(开d3)=1/31\*2=2/3

因此，此时应该改选d2 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313956#summary 作者：喵星人拯救地球

强烈参考 《三扇门的问题》 https://blog.csdn.net/ZCSYLJ/article/details/8113384

60. 分布函数具有可加性的性质很重要，以下说法正确的是？

(1) 均匀分布 (2) 二项分布 (3) 泊松分布 (4) 正态分布 (5) 伽马分布 (6) 卡方分布

A. 除了(1)，其它具有可加性 B. 除了(2)，其它具有可加性 C. 除了(3)，其它具有可加性 D. 都具有可加性

正确答案：A

「题目解析」 均匀分布不具有可加性。 图片来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313958#summary 作者：兰陵王＆阿兰朵

61. 京东Joy是个懂得分享的小朋友，他有10个相同的糖果，分给3个Jder，每个人至少要得到一个。有多少不同的发配方法？

A. 33 B. 34 C. 35 D. 36

正确答案：D

「题目解析」 用切分法。理解成长度为10的线段，有9个空档可以选择，切割两刀，分成三份。 C9(2) = 36。

这道题和《网易2018校园招聘数据分析工程师笔试卷》的第11题思路一致，举一反三。

62. 京东图书开展赠书活动，采取一种排列规则。有20名Jder有幸被选中参加赠书活动。20名Jder配安排站为4排，每排5个人。从中任选4Jder赠送一本最新的图书，那么我们选的人都在不同排的概率为（）

A. 5^4 \* 5!15!/20! B. 4^5 \* 5!15!/20! C. 5^4 \* 4!16!/20! D. 4^5 \* 4!16!/20!

正确答案：C

「民间解析」

不考虑排队，20个人20个坑，一共有20！种放法，从4个队列中拿出4个坑，一共有 5^4种方法，再把四个人放到4个坑中有4!种放法，剩下16个人放在16个坑中有16！种放法 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313960#summary 作者：Charles201809301636386

63. 京东Joy有有若干5g和7g的砝码，任何大于（ ）克都能够用5g和7g的砝码组合出。

A. 35 B. 23 C. 12 D. 53

正确答案：B

「民间解析」

解法1

假设ag可以用5g和7g的砝码组合出，且a后的所有值都可以用5g和7g的砝码组合出。 记：a=5m+7n (m,n为自然数） 我们来看看m和n的最小值！也就是说看看a中至少要含有多少个5多少个7。 讨论：设k为正整数！ 要能表示a+1，那么n>=2 (至少需要用2个7换成3个5） 要能表示a+2，那么m>=1 (至少需要用1个5换成1个7） 要能表示a+3，那么n>=1 (至少需要用1个7换成2个5） 要能表示a+4，那么m>=2 (至少需要用2个5换成2个7） 要能表示a+5，直接＋5就可以了！ 以后就有： a+6=(a+1)+5 a+7=(a+2)+5… 综合上述讨论知道：m>=2 n>=2 a=5m+7n>=2\*(5+7)=24>23 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313961#summary 作者：weixy18

解法2：

可以试想 假设满足条件最小的数是n，本身可以拆成若干5和若干7 ，那么比他大1 n+1 也可以拆成5和7 ，那么如何加1呢 最少要 加3个五克砝码 减2个七克砝码 类似的 如果要n+2也满足 至少需要 加1个7 减1个5 n+3 需要 加2个5 减1个7 n+4 需奥 加2个7 减2个5 n+5 以上，只需要增加5和7砝码的数量即可 考虑所有需要减少的砝码 对于n 至少要能拆成2个5 和2个7 才能实现+1 +2 +3 +4的操作 所以n至少是24 也就是25 26 27…都可以 题目问大于几 自然是23了 选B 来源：https://www.nowcoder.com/test/question/done?tid=24816459&qid=313961#summary 作者：weixy18

解法3

互质的两数p,q ，则重量为pq-p-q的砝码是不能被称出的最重的砝码 pq-p-q=(p-1)q-p，既不是p的倍数也不是q的倍数 对于原题，5\*7-5-7=23即为答案 来源：http://page.renren.com/601505826/note/877505552

举一反三 《网易2018校园招聘数据分析工程师笔试卷》第13题