

阿里巴巴集团 2014 年秋季校园招聘

系统工程师 北京 笔试卷

第一部分 单选题（前 10 题，每题 2 分；后 10 题，每题 3 分；共 50 分。选对得满分，选错倒扣一分，不选得 0 分。）

1. 字符串“alibaba”有_____个不同的排列。

A. 5040

B. 840

C. 14

D. 420

2. 下列一段 C++代码的输出是_____。

```
class Base
{
public:
    int Bar(char x) {return (int)(x);}
    virtual int Bar(int x) {return(2*x);}
};

class Derived :public Base
{
public:
    int Bar(char x) {return(int)(-x);}
    int Bar(int x) {return (x/2);}
};

void main(void)
{
    Derived Obj;
    Base*pObj=&Obj;
    printf("%d",pObj->Bar((char)(100)));
    printf("%d",pObj->Bar(100));
}
```

A. 100, 50
200

B. -100, 200

C. -100,50

D. 100,

3. 有一个二维数组 A[10][5]，每个数据元素占 1 个字节，且 A[0][0]的存储地址是 1000，则 A[i][j]的地址是_____。

A. $1000+10i+j$
1000+10i+5j

B. $1000+i+j$

C. $1000+5i+j$

D.

4. 下列_____不是线性表？

A. 队列
链表

B. 栈

C. 关联数组

D.

5. 下列有关在一个处理器（processor）上跑两个线程（thread）的说法中，正确的是_____。

- A. 一个线程可以改变另一个线程的程序计数器（program counter）
- B. 一个线程既不能读也不能写另一个线程的栈（stack）
- C. 一个线程可以读写另一个线程的寄存器（register）
- D. 以上都不对

6. 关于双链表的搜索给定元素操作的正确的说法是_____。

- A. 从两个方向搜索双链表，比从一个方向搜索双链表的速度慢
- B. 从两个方向搜索双链表，比从一个方向搜索双链表的方差要小
- C. 从两个方向搜索双链表，比从一个方向搜索双链表速度要快
- D. 以上说法都不正确

7. 对 n 个数字进行排序，期中两两不同的数字的个数为 k ， n 远远大于 k ，而 n 的取值区间长度超过了内存的大小，时间复杂度最小可以是_____。

- A. $O(n \log k)$
- B. $O(nk)$
- C. $O(n)$
- D. $O(n \log n)$

8. 一台指针式钟表的时钟和分钟的指向重合的时间间隔是_____ B_____。

- A. 720/13 分钟
- B. 720/11 分钟
- C. 60 分钟
- D. 以上都不正确

9. 两个大小不同的杯子 R 和 S，R 中装着一定量的小米，S 中装着一定量的沙子。一名儿童用勺子从 S 中取出一勺沙子放入 R，与小米混合之后，再从 R 中取出等体积的一勺混合物放入 S。假定两勺物品的体积相等，且 R 和 S 都没有发生溢出。则以下说法中正确的是_____。

- A. R 中的沙子和 S 中的小米一样多
- B. R 中的沙子比 S 中的小米少
- C. R 中的沙子比 S 中的小米多
- D. 无法判断

10. 假定抛出的硬币落地之后正反两面出现的概率分别是 $1/2$ 。那么抛 10 次和 100 次硬币（分别称为 T10 和 T100）相比，以下说法正确的是_____。

- A. T100 出现一半的正面比 T10 出现一半正面的概率更大
- B. T100 前 3 次都是正面的概率比 T10 前 3 次都是正面的可能性大
- C. T100 正面次数的方差小于 T10 出现正面次数的方差
- D. T100 出现正面的比例比 T10 出现正面的比例在 $(0.45, 0.55)$ 区间中的可能性更大。

11. 某福彩机构推出了一款简单的猜谜游戏：游戏玩家只需交纳 n 元，赌红或者黑。如果开奖结果与游戏玩家所赌的颜色相同，则玩家除得到交纳的 n 元赌资外，还可以获得 n 元作为奖励；否则该玩家失去交纳的 n 元赌资。为了游戏公平，开奖是红或者黑的概率均为 $1/2$ 。某游戏玩家想出了一个玩法：开始出 100 元参与赌博，然后按照如下规则进行游戏，如果输掉，并且赌资充足，就把已经输了的总钱数翻倍作为赌资进行赌博；否则，就停止该游戏。假定该机构赌资无限，而玩家的赌资比较有限，以下关于该玩家退出游戏时的情形的评论中合理的是：_____。

- A. 该玩家的策略可以保证游戏结束时赢钱数的期望为正数
- B. 该福彩机构长期会赔钱
- C. 该玩家会有一定概率在游戏结束时输钱，但输得不多
- D. 该玩家赢的可能性比输的可能性大

12. 有 16 瓶水，其中只有一瓶水有毒，小白鼠喝一滴之后一小时会死。请问最少用_____只小白鼠，在 1 小时内一定可以找出至少 14 瓶无毒的水？

- A. 1 只
- B. 3 只
- C. 4 只
- D. 16 只

13. 有一台 4 核 CPU 的服务器，上面运行着 1 种在线服务。如果该在线服务处理一个请求在非独占 IO 上的等待时间和 CPU 计算上消耗的时间比为 2:1，假定 IO 带宽充足，那么至少开_____个线程能最大化性能地使用该服务器？

- A. 4
- B. 8
- C. 12
- D. 线程越多越好

14. 有一种语言称为 lua，里面的数字只有一种类型（number），实际上是双精度浮点数。没有各种位数的整数，如 32 位、64 位整数等。那么关于该语言的说法错误的是_____。

- A. 该语言可以用 number 类型的变量作为数组下标
- B. 该语言可以表示任意 32 位数字整数的数字 ID
- C. 该语言无法实现 32 位数字整数的按位与、或、异或运算
- D. 该语言可以正常进行双精度浮点数运算

15. 一个在线服务通常需要读取存储着海量数据的数据库。为了提高服务的处理速度，通常需要加 cache（缓存）。以下场景中不合适使用 cache 的是_____。

- A. 数据库中每条数据被访问到的概率近似相等，且独立
- B. 使用了多线程机制的服务
- C. 单条数据尺寸大小的数据
- D. 有着大量访问的服务

16. 如下一段神奇的代码实现的功能是_____。

```
int miracle(unsigned int n)
{
    int m=n==0 ? 0:1;
    while (n=(n&(n-1)))
    {
        m++;
    }
    return m;
}
```

- A. n 的二进制表示中“0”的个数
- B. n 的二进制表示的倒序值
- C. n 的二进制表示中“1”的个数
- D. 一个均匀的哈希函数

17. 有 1023 个两两不同的整数，取值范围是 1 到 1024，其按位异或的结果的取值范围是_____。

- A. 0 到 1024
- B. 0,1025 到 2047
- C. 1 到 1024
- D. 2 到 2048 之前的全部偶数

18. 七夕节 n 恋人 ($n \geq 2$) 围成一圈举行篝火晚会。晚会的规则是：男女相同，且每对恋人处在相邻的位置上。请问有多少种不同的圈子？

- A. $(2n-1)!/2$
- B. $2(n-1)!$
- C. $2^n(n-1)!$
- D. $(2n)!$

19. 星期天有 10 个朋友约好一起郊游，在车站的集合时间是早晨 9:50:00 到 10:00:00。已知每个人到达车站的时间是 9:50:00 到 10:00:00 内的均匀分布，且彼此独立。那么最后一人最可能到达的时间是_____（精确到分钟，向下取整）。

- A. 各个分钟概率相等
- B. 9:57
- C. 9:58
- D. 9:59

20. 已知某国家每年出生人口数每年递增 3%，且男女比例为 1:1。如果每个男性都希望找比自己小 0.5 到 3 岁的女性结婚，且每个女性都希望找比自己大 0.5 到 3 岁的男性结婚，适婚年龄为 20 到 30 岁，那么对该国适婚男女婚配方面的说法正确的是_____。

- A. 男女会比较均衡
- B. 会产生较多剩女
- C. 会产生较多剩男
- D. 信息不足，无法判断

第二部分 不定向选项（4 题，每题 5 分。每题有 1-5 个正确选项，完全正确计 5 分，漏选计 2 分，不选计 0 分，多选、错选计-2 分）

21. 要提高多线程程序的效率，对锁的控制策略非常重要。一种策略是在锁的个数不太多、控制结构不太复杂的情况下，尽可能降低加锁的粒度；另一种策略是在合适的条件下取消用锁。以下情况中不可能取消锁的是_____。

- A. 多线程写一个共同的数据结构，且写操作是原子操作
- B. 多线程写一个共同的数据结构，且写操作不是原子操作
- C. 多线程读一个共同的数据结构，且读操作不是原子操作
- D. 一个线程写，多个线程读一个共同的数据结构，写操作是原子操作，读操作不是原子操作
- E. 一个线程写、多个线程读一个共同的数据结构，写操作不是原子操作，读操作是原子操作

22. 一颗非空的二叉树的先序遍历序列与后序遍历序列正好相反，则该二叉树可能是_____。

- A. 所有的结点均无右孩子
- B. 只有一个叶子结点
- C. 是一颗二叉树索树
- D. 所有的结点均无左孩子

23. 以下数字在表示为 double（8 字节的双精度浮点数）时存在舍入误差的有_____。

- A. $\sqrt{2}$
- B. 10 的 30 次方
- C. 0.1
- D. 0.5
- E. 100

24. 给定如下 C 程序：

```
typedef struct node_s{
int item;
struct node_s* next;
}node_t;

void reverse_list(node_t* head)
{
    node_t* n=head;
    head=NULL;
    while(n){
        _____
    }
    return head;
}
```

以下哪项能实现该函数的功能

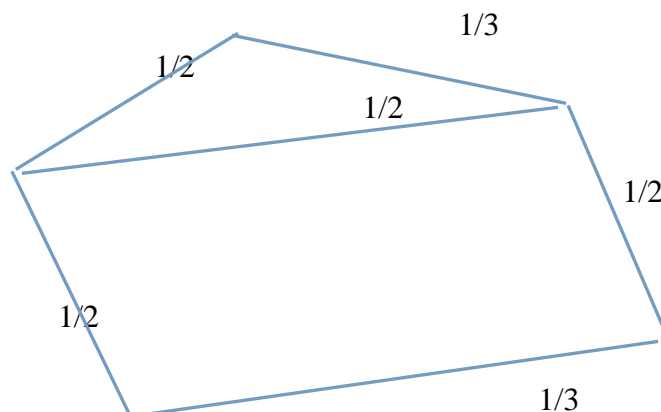
- A. `node_t* m=head; head=n; head->next=m; n=n->next;`
- B. `node_t* m=n; n=n->next; m->next=head; head=m;`
- C. `node_t* m=n->next; n->next=head; n=m; head=n;`
- D. `head=n->next; head->next=n; n=n->next;`

第三部分 填空与问答（5 题，共 30 分）

25.（4 分）某无聊的程序员在玩 Windows 上的记事本程序，不用鼠标，每次可以按以下键或组合之一：A、Ctrl+A（全选）、Ctrl+C（拷贝）、Ctrl+V（粘贴），那么在 10 次按键只能可以制造的最长文本长度为_____。

26.（4 分）若初始序列为 gbfcdae，那么只会少需要_____次两两交换，才能使该序列变为 abcdefg。任给一个自由 a--g 这 7 个字母组成的排列，最坏的情况下需要至少_____次两两交换，才能使序列变为 abcdefg。

27.（5 分）在某恶劣天气，若地图上 S 点到 T 点的交通网如下图所示，其中每条边表示一条双向通道，其上的数字为该通路可通行的概率，且该概率两两独立。求 S 到 T 的可通行概率_____。



28.（8 分）6 度分离假说的含义是，世界上任何两个人要么是朋友，要么是朋友的朋友，或者更高阶的朋友的朋友（如朋友的朋友的朋友），改论断中“朋友”一词出现的次数为两人之间的距离，那么该距离小于等于 6。如果某 SNS（如 QQ、旺旺等），有 100 万用户，其人际关系网略符合以下两个假设：

- 1) 朋友关系是一种对称关系（如 A 和 B 是朋友，那么 B 和 A 也是朋友）
- 2) 符合 2 度分离假说
- 3) 第 i 个人拥有的朋友的个数为 n_i ，所有 n_i 中最大值为 n

试估算 n 的最小值_____。

29. (9 分) 某电子商务网站进行 A、B 两种推荐算法的效果对比测试，对用户的访问请求按照 1:9 的比例随机分配给 A 和 B 两种算法处理。产生推荐结果后，按照两种指标对比两种算法产生的结果好坏：第一种指标是 CTRPV=该算法下用户的点击展现次数/该算法下所有的展现次数，第二种指标是 CTRUV=该算法下有点击的用户数/该算法下所有的用户数。假定每个用户会对该推荐服务 2 次访问，如果 A 和 B 的 CTRPV 持平（假设为 0.01）。那么 CTRUV 哪个大，大的比小的大百分之多少_____。

第四部分：JAVA 附加题（注，阿里有大量 JAVA 研发工程师需求；选作以下题目有机会增加该方向面试机会）

1. 以下每个线程输出的结果是什么？（不用关注输出的顺序，只需写出输出的结果集即可）

```
public class TestThread{

    public static void main(String[] args){
        // test1
        Thread t1 = new Thread(){
            @Override
            public void run(){
                try{
                    int i=0;
                    while(i++<100000000){
                        // nothing
                    }
                    System.out.println("A1");
                }catch(Exception e){
                    System.out.println("B1");
                }
            }
        };
        t1.start();
        t1.interrupt();

        // test2
        Thread t2 = new Thread(){
            public void run(){
                try{
                    Thread.sleep(5000);
                    System.out.println("A2");
                }catch(Exception e){
                    System.out.println("B2");
                }
            }
        };
    }
}
```

```

        };
    };
    t2.start();
    // t2.interrupt();    //不确定是否有这句话

    // test3
    Thread t3 = new Thread(){
        public void run(){
            try{
                Thread.sleep(50000);
                System.out.println("A3");
            }catch(Exception e){
                System.out.println("B3");
            }
        }
    };
    t3.start();
    t3.interrupt();

    // test4
    Thread t4 = new Thread(){
        public void run(){
            try{
                Thread.sleep(50000);
                System.out.println("A4");
            }catch(Exception e){
                System.out.println("B4");
            }
        }
    };
    t4.start();
    t4.interrupt();

    // test5
    try{
        t4.start();
        System.out.println("A5");
    }catch(){
        System.out.println("B5");
    }
}
}

```


2. 一个 10 亿条记录的文本文件，已按照关键字排好序存储，请设计方法，可以快速地从文件中查找指定关键字的记录。

阿里巴巴集团 2014 年秋季校园招聘

【系统工程师】 附加题

1.在互联网时代系统的稳定性要求越来越高，为了提升系统的稳定性，高可用技术被广泛运用，请列举至少 4 中相关的技术解决硬件、系统或网络等层面的单点问题。

2.请描述一下 TCP 建立连接三次握手的过程。

3. 搜索引擎是很常用的 web 应用。大部分搜索引擎需要设计一个抓虫(Crawler)，从很多网站抓去网页，分析数据，供搜索引擎使用。

设想你来做一个搜索引擎的爬虫，需要抓去约一百万家网站的网页内容。

- 1) 请画出一个抓虫系统的架构图。
- 2) 重点说明你的爬虫需要如何优化来提升性能。