选择题  
1、已知中国人的血型分布约为A型：30%，B型：20%，O型：40%，AB型：10%，则任选一批中国人作为用户调研对象，希望他们中至少有一个是B型血的可能性不低于90%，那么最少需要选多少人？  
A、7     B、9      C、11     D、13  
2、广告系统为了做地理位置定向，将IPV4分割为627672个区间，并标识了地理位置信息，区间之间无重叠，用二分查找将IP地址映射到地理位置信息，请问在最坏的情况下，需要查找多少次？  
A、17     B、18      C、19     D、20  
3、有四只老鼠一块出去偷食物（每个都偷了），回来时，族长问它们都偷了什么，老鼠A说：我们每个都偷了奶酪。老鼠B说：我只偷了一颗樱桃。老鼠C说：我没偷奶酪。老鼠D说：有些人没偷奶酪。族长观察了一下，发现它们当中只有一只老鼠说了实话，那么是哪只老鼠说了实话？  
A、老鼠A        B、老鼠B        C、老鼠C       D、老鼠D  
4、到商店里买200的商品返还100的优惠券（可以在本商店代替现金）。如果使用优惠券买东西不能获得新的优惠券，那么买200返100优惠券，实际上省多少？  
A、50%  
B、66.7%  
C、75%  
D、33.3%  
5、在数据库逻辑设计中，当将E-R图转换为关系模式时，下面的做法哪一个不正确？  
A、一个实体类型转换为一个关系模式  
B、一个联系类型转换为一个关系模式  
C、由实体类型转换成的关系模式的主键是该实体类型的主键  
D、由联系类型转换成的关系模式的属性是与该联系类型相关的诸实体类型的属性的全体  
6、一家人有两个孩子，性别未知，现在打电话给其中一个孩子得知是女孩，问另一个孩子也是女孩的概率是多少？  
A、1/4        B、1/2        C、1/3       D、1/5  
7、关于非空二叉树的性质，下面哪个结论不正确(D)  
A、有两个节点的节点一定比没有子节点的节点少一个      n0 = n2 + 1  
B、根节点所在的层数为第0层，则第i层最多有2^i个节点  
C、若知道二叉树的前序遍历序列和中序遍历序列，则一定可以推出后序遍历序列。  
D、堆一定是一个完全二叉树  
  
8、快速排序的平均时间复杂度和最坏时间复杂度是（）  
A、O（n^2）, O(n^2)  
B、O（n^2）, O(nlgn)  
C、O(nlgn) , O(nlgn)  
D、O(nlgn) , O(n^2)  
9、有一串数字 6 7 4 2 8 1 6 （），请问括号中的数字最可能是（）  
A、6       B、7       C、8       D、9  
10、下面哪项不是链表优于数组的特点？  
A、方便删除     B、方便插入    C、长度可变    D、存储空间小  
11、给定声明 const char \* const \* pp; 下属操作或说明正确的是（）  
A、pp++  
B、(\*pp)++  
C、(\*\*pp) = 'c';  
D、以上都不对  
12、有下列代码正确的是（）

**[cpp]** [view plaincopyprint?](http://blog.csdn.net/hackbuteer1/article/details/8017749)

1. std::string name1 = "youku";
2. const **char**\* name2 = "youku";
3. **char** name3[] = {'y','o','u','k','u'};
4. **size\_t** l1 = name1.size();
5. **size\_t** l2 = strlen(name2);
6. **size\_t** l3 = sizeof(name2);
7. **size\_t** l4 = sizeof(name3);
8. **size\_t** l5 = strlen(name3);

std::string name1 = "youku";

const char\* name2 = "youku";

char name3[] = {'y','o','u','k','u'};

size\_t l1 = name1.size();

size\_t l2 = strlen(name2);

size\_t l3 = sizeof(name2);

size\_t l4 = sizeof(name3);

size\_t l5 = strlen(name3);

A、l1 = 5     l2 = 5     l3 = 4      l4 = 5     l5 = 不确定  
B、l1 = 5     l2 = 5     l3 = 5      l4 = 5     l5 = 不确定  
C、l1 = 5     l2 = 6     l3 = 5      l4 = 5     l5 = 5  
D、l1 = 5     l2 = 6     l3 = 5      l4 = 5     l5 = 6  
13、Test执行后的输出是：

**[cpp]** [view plaincopyprint?](http://blog.csdn.net/hackbuteer1/article/details/8017749)

1. void Test()
2. {
3. class B
4. {
5. public:
6. B(void)
7. {
8. cout<<"B\t";
9. }
10. ~B(void)
11. {
12. cout<<"~B\t";
13. }
14. };
15. struct C
16. {
17. C(void)
18. {
19. cout<<"C\t";
20. }
21. ~C(void)
22. {
23. cout<<"~C\t";
24. }
25. };
26. struct D : B
27. {
28. D()
29. {
30. cout<<"D\t";
31. }
32. ~D()
33. {
34. cout<<"~D\t";
35. }
36. private:
37. C c;
38. };
39. D d;
40. }

void Test()

{

class B

{

public:

B(void)

{

cout<<"B\t";

}

~B(void)

{

cout<<"~B\t";

}

};

struct C

{

C(void)

{

cout<<"C\t";

}

~C(void)

{

cout<<"~C\t";

}

};

struct D : B

{

D()

{

cout<<"D\t";

}

~D()

{

cout<<"~D\t";

}

private:

C c;

};

D d;

}

A、B   C   D   ~D   ~ C   ~B  
B、D   C   B   ~B   ~C   ~D  
C、C   D   B   ~B   ~D   ~C  
D、C   ~C   D   ~D   B   ~B  
14、下列四种排序中（D）的空间复杂度最大  
A、快速排序        B、冒泡排序       C、希尔排序         D、堆  
15、设一棵二叉树的深度为k，则该二叉树最多有（D）个节点。  
A、2k-1      B、2^k    C、2^(k-1)      D、2^k-1  
16、下面函数的功能是（）

**[cpp]** [view plaincopyprint?](http://blog.csdn.net/hackbuteer1/article/details/8017749)

1. **int** fun(**char** \*x)
2. {
3. **char** \*y = x;
4. while(\*y++);
5. return (y-x-1);
6. }

int fun(char \*x)

{

char \*y = x;

while(\*y++);

return (y-x-1);

}

A、求字符串的长度  
B、比较两个字符串的大小  
C、将字符串x复制到字符串y  
D、将字符串x连接到字符串y后面  
17、k为int类型，以下while循环执行（）次。

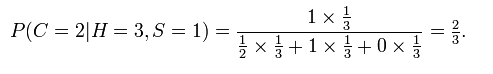
**[cpp]** [view plaincopyprint?](http://blog.csdn.net/hackbuteer1/article/details/8017749)

1. unsigned **int**  k = 20;
2. while(k >= 0)
3. --k;

unsigned int k = 20;

while(k >= 0)

--k;

A、20次    B、一次也不执行   C、死循环   D、21次  
18、关于Cookie 和 Session的概念哪一个是对的  
A、Cookie 存储在客户端，但过期时间设置在服务器上  
B、Session 存储在客户端，但过期时间设置在服务器上  
C、Cookie 中可以存储ASCII空格‘  ’，而Session中不行  
D、Cookie可以设置生效的路径，而 Session则不能  
19、以下关于链式存储结构的叙述中哪一条是不正确的？  
A、结点除自身信息外还包括指针域，因此存储密度小于顺序存储结构  
B、逻辑上相邻的结点物理上不必邻接  
C、可以通过计算直接确定第i个结点的存储地址  
D、插入、删除运算操作方便，不必移动结点  
20、32位机器上，定义 int \*\*a[3][4]，这个数组占多大的空间（）  
A、64          B、12        C、48        D、128  
  
填空题  
1、设数组定义为a[60][70]，每个元素占2个存储单元，数组按照列优先存储，元素a[0][0]的地址为1024，那么元素a[32][58]的地址为（8048）  
2、在一个娱乐节目上，主持人提供有三扇门（假设为A、B、C），只有1扇门后面有奖品，另两扇门后面是空的，而主持人知道具体哪扇门后有奖品。首先，当你选择了一扇门之后（假设A），主持人会把剩下两扇门中的一扇没有奖品的门打开（假设打开的空门为B），现在你有一次机会决定是否要交换重新选择，如果你坚持选择A，你中奖的概率是（1/3），如果你交换选择C，你中奖的概率是（2/3）   <http://en.wikipedia.org/wiki/Monty_Hall_problem>  
假设你选择的1门，而主持人打开的是3门，则奖品在2门后面的概率是  
  
  
3、一棵深度为h的满二叉树，其最末一层共有（2^h）个节点（根节点深度为0）  
4、下面程序的运行结果为（1   3   2）

**[cpp]** [view plaincopyprint?](http://blog.csdn.net/hackbuteer1/article/details/8017749)

1. void foo(**int** \*a , **int** \*b)
2. {
3. \*a = \*a + \*b;
4. \*b = \*a - \*b;
5. \*a = \*a - \*b;
6. }
8. void main()
9. {
10. **int** a = 1 , b = 2 , c = 3;
11. foo(&a , &b);
12. foo(&b , &c);
13. foo(&c , &a);
14. printf("%d %d %d\n",a,b,c);
15. }

void foo(int \*a , int \*b)

{

\*a = \*a + \*b;

\*b = \*a - \*b;

\*a = \*a - \*b;

}

void main()

{

int a = 1 , b = 2 , c = 3;

foo(&a , &b);

foo(&b , &c);

foo(&c , &a);

printf("%d %d %d\n",a,b,c);

}

5、4个结点可以构造出（14）个不同的二叉树    Catalan数  
6、设有n个无序的记录关键字，则直接插入排序的时间复杂度为（O(n^2)），快速排序的平均时间复杂度为（O(nlgn)）  
7、设一组初始记录关键字序列为（20,18,22,16,30,19），则以20为中轴的一趟快速排序结果为（19,18,16,20,30,22）  
8、C语言的函数参数传递方式有传递 值   和  传递 地址  
9、分配在堆上和栈上的内存，哪一个需要手动进行内存释放？   堆上的内存  
  
问答题：  
一、有一个单向循环链表队列，从头开始报数，当报到m或者m的倍数的元素出列，根据出列的先后顺序重新组成单向循环链表。  
函数原型： void reorder(Node \*\*head , int m)  
  
二、优酷是中国第一的视频网站，每天有上亿的视频被观看，现在公司请研发人员找出最热门的视频。  
该问题的输入可以简化为一个字符串文件，每一行都表示一个视频id，然后要找出出现次数最多的前100个视频id，将其输出，同时输出该视频的出现次数。  
1、假设每天的视频播放次数为3亿次，被观看的视频数量为一百万个，每个视频ID的长度为20个字节，限定使用的内存为1G。请先描述做法，再写代码。  
2、假设每个月的视频播放次数为100亿次，被观看的视频数量为1亿个，每个视频ID的长度为20个字节，一台机器被限定使用的内存为1G。  
那么想找这个月被播放次数最多的前100个视频，应该怎么做？请描述清楚可能的办法。  
解析：海量数据的处理。无法一次性装入内存，可先hash之分为多个文件处理，堆或者Trie树统计次数，求出每个文件中的Top 100。归并之求出总的top 100。  
  对于第二问：还可以hadoop mapReduce处理之。  
  首先统计每个视频被观看次数，得到<id, cnt>键值对，其中id为视频id，cnt为视频被观看次数。   
  以cnt作为关键字建立最小堆。遍历所有键值对，若堆的size小于100，则将键值对直接插入堆，否则比较键值对和堆顶元素大小，若cnt大于堆顶元素的cnt，则弹  出堆顶元素并将键值对插入堆。  
  对于第一问，由于id个数较少，统计部分可直接使用stl的map容器。   
  对于第二问，由于id个数太大，直接hash内存不够，需要mapReduce。  
三、给你一个由n-1个整数组成的未排序的序列，其元素都是1到n中的不同的整数。请写出一个寻找序列中缺失整数的线性时间算法。