技术类－笔试题 笔试时间：90分钟

姓名：xxx

笔试题说明：

1. 从所有题目中任选4题作答，60分以上进入面试。正确回答更多的题目会有更高的笔试成绩。高笔试成绩者优先。
2. 所有题目必须从C、C++、Java、Scala、Python、Go六种语言中选择一种实现
3. 请选择尽可能高效的算法，算法复杂度（大O表示法）比最优解高的话，得一半分。

**题目一.** 有两个有序数组，合并为一个有序数组。（15分）

示例：

输入：

arrayA : [1， 3， 5， 7， 9]

arrayB : [2, 4, 6, 8, 10]

输出：

[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

**题目二.** 有文本文件text.txt，文件中每一行代表一行文本内容，现需要统计某词列表keyword.lst中的词的命中次数（即一行文本中包含该词），除此之外，需对每个词所命中的文本进行保存，命中次数大于100的保留100条，对命中次数小于100的保留全部文本。（25）

示例：

输入：

text.txt：

hello world, Java

Have a good day

Nice to meet you

hello world, C++

keyword.lst：

hello

world

day

输出：

统计：

hello: 2

world: 2

day: 1

文本详情：

hello:

hello world, Java

hello world, C++

world:

hello world, Java

hello world, C++

day:

Have a good day

**题目三.** 把中文类型的数字表示形式转换成数字串，类似“伍拾叁”转换为“53”，“玖拾捌亿柒仟陆佰伍拾肆万叁仟贰佰零壹”转换为“9876543201”。（30分）

示例：

输入（从标准输入读取）：

玖拾捌亿柒仟陆佰伍拾肆万叁仟贰佰零壹

输出（打印到标准输出）：

9876543201

**题目四.** “特殊数字”是指只能被分解为2，3，7的乘积的数字。“特殊数字”序列为1, 2, 3, 4, 6, 7, 8, 9, 12, 14, ...展示了前10个“特殊数字”。按照惯例，1也是“特殊数字”。给定整数n，写一个程序，找到序列中，第n个“特殊数字”。如：n=9，找到“特殊数字”序列中第9位数字，也就是输出12。（30分）

输入：

每一行包含一个正整数，n (n <= 1500)。如果输入的一行为0，则输入终止。

输出：

对于每一行输入，输出“特殊数字”。注意不要解析输入为0的行

示例：

输入：

1

2

9

0

输出：

1

2

12

**题目五.** 有一个模型可以把一个串中关键字（微信，QQ等）和联系方式的串识别出来，例如输入一段文字“想知道如果提高王者荣耀水平，上分把妹不是梦，加微信：17252sugats78，加QQ：34676328889”，则模型的输出结果为每个自然字对应的标签（1表示关键字开头，2表示关键字的中间和末尾，3表示串的开头，4表示串的中间和结尾，0表示无关变量），例如上面那段文字的输出如下[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,2,0,3,4,4,4,4,4,4,4,4,4,4,4,4,0,0,1,2,0,3,4,4,4,4,4,4,4,4,4,4]  
请写一段程序根据输入的文本和模型给出的预测标签输出格式化结果。（30分）

示例：

输入文本：  
想知道如果提高王者荣耀水平，上分把妹不是梦，加微信：17252sugats78，加QQ：34676328889

输入标签：

[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,2,0,3,4,4,4,4,4,4,4,4,4,4,4,4,0,0,1,2,0,3,4,4,4,4,4,4,4,4,4,4]

输出格式化文本：  
关键字：['微信','QQ']，串:['17252sugats78','34676328889']