

# 2020 年夏季 Java 小学期

## Homework5

### 问答题

#### 【约定】

- 本题要求将答案写到答题卡.xlsx 的作答区域中。
  - 问题保证答题卡.xlsx 文档没有使用任何宏。
- 提交的答卷格式为.xlsx。如果不支持此格式，可以将文件转为答题卡.xls 或答题卡.csv。
  - 如果文件被保存为.csv 格式，请确保编码被设置为 UTF-8。但不建议在.csv 格式下作答，因为会丢失选择题的数据验证。
- 提交结果为半自动批阅，请不要私自修改非作答区域的内容，否则可能出现答案无法读取。
  - 可以修改行宽、列宽、字体、字号，但不要增加 sheet 或改变答案的行号、列号，也不可以合并单元格。
  - 建议直接使用默认字体，而不要使用自定义字体，自定义字体可能使用特殊编码导致内容无法读取。
- 回答多选题时请从左到右按照编号顺序填写选项，一个区域一个选项，多余区域保留空白选项。

#### 【第一题】选择

- (1) 在 Java 中，引入了对象互斥锁的概念，用什么关键字可以令当前线程获得对象的互斥锁？
- A. try
  - B. pthread
  - C. parallel
  - D. synchronized
  - E. return
- (2) 以下说法正确的是？
- A. 一个线程可以同时持有多个对象的互斥锁
  - B. 在包 java.util 中，容器 Vector, Hashtable 都是并发安全的容器，而 ArrayList, HashMap 都是非并发安全的容器
  - C. 如果有多个线程需要打印同一个 String 对象，则必须使用互斥锁，否则各个线程会因为并发错误打出不一样的字符串
  - D. 对于变量 int a，操作 a++ 是并发安全的，不必依赖互斥锁
  - E. Thread 的 sleep() 方法会令当前线程睡眠，并释放已持有的锁资源，让其他线程获得使用权
- (3) 以下说法正确的是？
- A. 已经停止的线程，可以通过调用其 start() 方法令其重新运行
  - B. 如果要调用对象 obj 的 wait()、notify() 或 notifyAll() 方法，必须先持有对象 obj 的互斥锁

斥锁

- C. Java 的线程优先级为从 1 到 10 的整数，10 表示最大的优先级
- D. UDP 协议以数据包的形式传输数据，而 TCP 协议提供流式数据传输，UDP 是不可靠的，TCP 是可靠的
- E. TCP 协议需要建立连接，建立连接的方法是两端一起使用 Socket 的 connect 方法向对方请求连接

(4) 以下说法正确的是？

- A. 想获取一个 URL 资源的内容，可以调用它的 openConnection()方法
- B. 要获取百度的网页内容，可以直接调用 new URL("baidu.com").openConnection()
- C. 在同一台主机上，可以同时通讯的 TCP 套接字最多只有 65536 个，对应不同的端口号
- D. 在同一台主机上，本质不同的套接字总共有 65536\*2 个，对应 65536 个端口号与 TCP 或 UDP 两种协议
- E. 在同一台主机上，两个不同的进程可以同时使用相同的套接字，因为分属不同进程，不会有资源冲突

## 编程题

【约定】

时间限制：5 s。

空间限制：1 GB。（VM 参数：-Xmx1GB）

程序入口类为 Main。

【第一题】机器哲学家

问题

某偏僻城市聚集着大量机器哲学家 MachinePhilosopher，哲学家有 N 个，编号 0~N-1。他们自己建立了一个公共频道，而且他们很喜欢查看这个频道上的各种消息。哲学家们在生活中的交流事件可以用如下模型描述。

有 1 个公共频道 Com，实例可以用 Com.getInstance()获得，哲学家首先要取得这个实例的互斥锁，然后开始他们各自的生活。

哲学家们的生活由方法 run()定义，这个生活是一个基于 Com 的实例构建的循环：

- 如果是初次执行循环，哲学家会调用 Com 实例的 getNumber()得到当前的消息编号，然后用 Com 实例的 wait()方法等待消息的更新
- 哲学家被唤醒后，再次调用 Com 实例的 getNumber()检查消息编号是否已更新，如果编号没有更新，回到循环开始
- 如果编号已更新，哲学家会更新自己记录下的消息编号，然后调用 Com 实例的 star(number, userId)方法表示在看，其中 number 对应的实参是当前消息的编号，userId 对应的实参是哲学家自己的编号

题目要求

为了减轻同学的工作量，本题提供一份包装好的代码，并给出了需要填充代码的类 MachinePhilosopher。

同学只需要实现 MachinePhilosopher 的方法 run()即可。

## 规模

哲学家的数量不超过 10 个。

消息总数不超过 1000 个。

## 附加代码

src/Com.java

src/MachinePhilosopher\_template.java

src/Main.java

## 提交文件

src/MachinePhilosopher.java

## 【第二题】关键词识别

### 问题

有一个已知的关键词列表，每个关键词是一个词或短语。

打开一个网页，从网页的代码里识别指定的关键词。

### 输入

第一行是需要打开的 URL。

后面每一行包括一个字符串，为要识别的关键词。

### 输出

若干行，每行一个在网页里出现的关键词，以及关键词的出现次数。优先按关键词的  
次数排序，次数相同时，按输入的顺序排序。

### 例子

#### 例 1

输入：

```
http://www.sample.com/  
如果  
果汁  
不如
```

某站 http://www.sample.com/:

```
牛奶不如果汁营养多
```

输出：

```
如果 1  
果汁 1  
不如 1
```

### 提示

看到这里，是不是很多同学疑惑如何找到站点？甚至发现本机 java 程序在打开 https://www.baidu.com/时会报错——证书认证不通过？这些都不重要，因为本题专门提供了网络代理，同学们只需要在解析 URL 请求之前运行如下语句即可：

```
URL.setURLStreamHandlerFactory(TUProxy::new)
```

使用代理类后，在运行时，要确保将附件中的 proxy 文件夹放到当前目录下。

另外，网页编码与输入文件的编码都是 UTF-8，在本地测试时，可以用如下方法确保正确解码：

```
java -Dfile.encoding=UTF-8 Main
```

### 附迷之 Bug

在 Windows 上的 Oracle Java 有可能在调用 new URL(...)构造器时花费大量时间（绝不少于 10 秒），如果你遇到这个情况不必慌张，此题最终测试会使用 Linux 服务器。

### 规模

实体数量不超过 100 个。

网站的文本长度不超过 10000 个字符，编码为 UTF-8。

实体的频率与阈值 threshold 都是 0~100000 的整数。

### 附加代码

src/TUProxy.java

### 提交文件

src/Main.java

### 【第三题】网络攻击

#### 问题

有一个类 Cloud，它的内部存储着宝贵的机密数据，但作为 Hacker 可以直接读取这些数据。Cloud 还会启动一个 ServerSocket，其监听的端口号为 11111，每当别的 Socket 与它建立连接并发送消息，它会进行如下操作。

- 从套接字读入第一行输入数据当作 message，然后关闭当前套接字
- 调用 Cloud.test(Integer.parseInt(message)) **【test 方法定义见附加代码】**
- 如果调用结果为 true，就打印“Cloud is hacked, loss data \*\*\*!”，然后更新被破解的数据
- 如果调用结果为 false 或抛出了异常，就打印“Game is over!”，然后停止服务
- 当 Cloud 被破解一定次数后，Cloud 会打印“Hacker wins!”，然后停止服务

你的任务是在类 Hacker 中实现函数 run()，让 Cloud 持续不断地被破解，直到 Cloud 宣布 Hacker 胜利。

#### 提示

请使用 InetAddress.getLoopbackAddress() 作为 Socket 的 IP 地址。

### 规模

需要的攻击次数不超过 1000 次。

### 附加代码

src/Cloud.java

src/Hacker\_template.java  
src/Main.java

### 提交文件

src/Hacker.java

## 【提交方法】

程序入口类名称为 Main，包名为默认，并放入相应的文件夹内。将解题过程写入实验报告中。关于程序的其他说明写入 readme.txt 中。并将以上文件打包为 zip 文件。提交至网络学堂。

文件打包如下：

zip 归档文件（命名为：学号\_姓名\_homework5.zip）

-- 答题卡.xlsx(.xls,.csv)

-- 实验报告.pdf（解题过程）

-- programming-1（编程题 1）

    -- src（目录内为提交的程序代码）

    -- readme.txt（可选，建议写）

-- programming-2（编程题 2）

    -- src（目录内为提交的程序代码）

    -- readme.txt（可选，建议写）

.....

## 【注意事项】

- 1、代码要求自己写。代码提交后会进行查重，如果被认定为抄袭，课程将会被记为 0 分。
- 2、按时提交。如有迟交，作业成绩在原有作业成绩基础上每迟交一天（不到一天按一天计）扣 10%。
- 3、题目给分结果由最终测试决定。

不要相信简单的附加代码与样例结果，最终测试时，附加代码，输入、输出文件都会变得更严格。