

苏州大学

2011 年硕士研究生入学考试初试试题 (A 卷)

科目代码: 853 科目名称: 专业综合(2) 满分: 150 分

注意: ①认真阅读答题纸上的注意事项; ②所有答案必须写在答题纸上, 写在本试题纸或草稿纸上均无效; ③本试题纸须随答题纸一起装入试题袋中交回!

生物信息学 部分 (150分)

本试卷共3大部分, 请选择其一作答。

一. 软件说明 (每题 8 分, 共 80 分)

1. MRBAYES
2. Haploview
3. TREECON
4. Jalview
5. TopPred II
6. Artemis
7. Jpred
8. HMMTOP
9. affy
10. GOstat

二. 名词解释 (每题 10 分, 共 50 分)

1. Gene Ontology
2. maximum likelihood
3. false discovery rate
4. statistical significance
5. Missing value in DNA microarrays

二. 问答题 (共 20 分)

1. Multiple sequence alignments: A) 说明其基本原理、常用方法 B) 如何提高精度?

Perl 编程部分 (150分)

一、判断对错。如果有错, 请解释理由。(3 x 10=30 分)

- a) 在 print 语句中使用换码序列\n 时, 会促使光标定位到屏幕下一行的起始处。
- b) 求余运算符(%)计算第一个运算对象除以第二个运算对象之后得到的余数。
- c) 如果 Perl 程序需要打印三行输出, 那么必须使用 3 个 print 语句。
- d) 数组里可保存多种不同类型的值。
- e) 数组下标可采用浮点数。
- f) \$hash[\$key]表达式可返回%hash 中的\$key 对应的值。
- g) print "%hash\n";语句会显示一个%hash 字样, 然后接上一个换行符。
- h) 如在 while、for 和 foreach 结构中执行 redo 语句, 循环会重新开始, 同时会对是否继续循环下去的条件进行检测。
- i) 子程序的主体只有在响应一个子程序调用时才会执行。
- j) 所有变量都要求插入包的符号表中。

二、指出下述语句的错误并改正:(4 x 5=20 分)

- a) print "The value is \$number. \n";
- b) if (\$c=>7) {print "c is equal to less than 7\n";}
- c) for(\$i=0, \$i<@numbers, ++\$i){print "\$numbers[\$i]";}
- d) @hash['key1','key2','key3']; #create hash slice
- e) \$string =~ m/[1-9];

三、请写出一些语句, 完成下述要求:(5 x 4=20 分)

- a) 将标量变量\$string 中的所有字符变为小写, 使用字符串函数和 tr//两种方法实现。
- b) 用 substr 函数将字符串"How are you?"(在标量变量\$string 中)更改为"How are thing?"。
- c) 将字符串"Is it sunny today" (在标量变量\$string 中)分解为独立的单词。
- d) 创建一个名为*new 的类型通配, 使其标量和散列表值成为*old 中标量和散列表的别名。但不能改变*new 其他类型的值。

四、写一个程序, 完成下述任务:(10 分)

- a) 创建两个 foreach 循环, 一个标签名为 PROMPT, 另一个标签名为 PRINTING。二者都是 foreach 循环, 而且 PRINTING 循环要嵌套在 PROMPT 循环内部。每个结构都在 1~10 的整数之间循环。(2 分)
- b) 在 PROMPT(外层)循环中, 提示用户输入一个值, 然后把这个值保存在标量变量 \$stars 中。(2 分)
- c) 在 PRINTING 循环中, 假如\$_大于\$stars 的值, 就打印一个换行符, 同时让程序立即进入下一个 PROMPT 循环; 否则的话, 打印一个星号, 并在后面跟一个空格。(6 分)
- d) 在每一组星号后面, 都打印一个换行符。(2 分)

五、假设有一段 HTML 文本保存在标量变量\$string 中, 写一程序, 用正则表达式挑出其中所有合法的 URL 地址。URL 的特征是: 它放在一对双引号内, 而且以"http://"开头。输出的 URL 要求去除双引号。(10 分)

六、写一个程序，能重复要求用户猜测某个在 1 到 100 之间的数字，直到猜对为止。程序应当能随机的产生一个数字，使用公式 $\text{int}(1+\text{rand}(100))$ 。当用户猜测错误时，程序应该回应 “Too high” 或者 “Too low”。如果用户输入 quit 或 exit，或者回车时，程序应立即退出。如果用户猜测正确，程序也退出。(15 分)

七、写一个程序，将下列散列表数据根据姓 (last name) 按照大小写无关的字母顺序进行排序，并把结果打印出来。当 last name 相同时，再按照名(first name)排序 (不用关心大小写)。因此，第一个输出的名字是 fred flintstone，最后一个名字是 betty rubble。具有相同 familyname 名字在一起。不要改变数据。输出名字的大小写应当和这里的一样。(15 分)

```
my %last_name=qw{
fred flintstone Wilma Flintstone Barney Rubble
betty rubble Bamm-Bamm Rubble PEBBLES FLINTSTONE
};
```

八、写一个程序在已有的 Perl 程序加上下面这样一行：(15 分)

```
##Copyright(C) 2009 by Yours Truly
```

将上面一行放在 “shebang” 行 (Perl 程序的第一行，如 `#!/usr/bin/perl`。注意：应当在 “源文件” 中修改，但请备份文件。假定你可以在命令行中同时输入程序和需要的修改的文件名。

九、写一个程序，从文件中读入字符串，一行一个字符串，然后让用户输入模式，这个模式可能匹配上某些字符串。对于每一个模式，程序将指出文件中有多少个字符串 (多少行) 匹配上了，并指出是哪些。对于新的模式不需要重新读文件，将这些字符串保留在内存中。文件名可以直接写在程序之中。如果模式无效 (例如，圆括号不匹配)，则程序报告这个错误，并让用户继续尝试新的模式。当用户输入一个空行，则程序退出。(15 分)

线形代数试卷部分 (150 分)

一、对矩阵 $A = \begin{pmatrix} 2 & 2 & 3 \\ 4 & 7 & 7 \\ -2 & 4 & 5 \end{pmatrix}$

(1) 求矩阵 A 的迹，行列式和逆矩阵；(15 分)

求解线性方程组： $Ax=b$ ，其中 $b=(7,18,7)^T$ 。(10 分)

二、 n 阶矩阵 $A = \begin{pmatrix} 1 & 3 & 3 & 3 & \cdots & 3 \\ 3 & 2 & 3 & 3 & \cdots & 3 \\ 3 & 3 & 3 & 3 & \cdots & 3 \\ 3 & 3 & 3 & 4 & \cdots & 3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 3 & 3 & 3 & 3 & \cdots & n \end{pmatrix}$

(1) 求 A 的行列式；(15 分)

(2) 判断 A 是否是正定矩阵，并说明理由。(10 分)：

三、对矩阵 $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}$

(1) 求 A 所有的特征值和特征向量；(10 分)

(2) 将矩阵 A 对角化，即求可逆矩阵 P ，使得 $P^{-1}AP$ 为对角矩阵；(10 分)

(3) 计算 A^{10} 。(5 分)

四、 A 是正交矩阵，证明

(1) A 的行列式等于 1 或 -1；(5 分)

(2) A 的特征值等于 1 或 -1。(10 分)

五、矩阵 $A=(a_{ij})_{n \times n}$ 是非奇异的下三角矩阵，即 $a_{ij}=0$ 当 $j>i$ 且 $|A| \neq 0$ ，证明

A 的逆矩阵仍然是下三角矩阵。(15 分)

六、矩阵 A 是 $m \times n$ 的实矩阵，证明

(1) $A^T A$ 是对称半正定矩阵；(10 分)

(2) 当 A 列满秩时， $A^T A$ 是对称正定矩阵。(15 分)

七、 A, B 都为对称正定矩阵，证明 AB 是对称正定矩阵的充分必要条件是 $AB=BA$ 。(20 分)