

阿里巴巴（中国）网络技术有限公司

2009 年校园招聘笔试题

试卷 B

答题说明：

- ◇ 请将您的答案写在答题纸上，并在答题纸顶端注明您的姓名、学校、专业及联系电话；
- ◇ 本试题分为卷一、卷二、卷三；卷一针对的是 Java 开发方面；卷二是针对搜索开发和 C++ 方向；请在答题前先选择应聘职位的方向；卷三为公共题，所有候选人都需要完成；
- ◇ 笔试答题时间为 60 分钟；
- ◇ 答完题后将考题及答案交至主考官处，请勿将试卷带出考场；
- ◇ 公司在每个笔试考点均安排有专业技术人员和人力资源部代表，如果您有何疑问请及时向他们咨询；
- ◇ 如果您通过了笔试，我们将随后的一周时间内通知您参加面试，具体时间和地点将电话（邮件）通知，请您届时保持联系方式的畅通。

B2B-技术部-工程师笔试试题 (卷一)

卷一: Java 开发、测试工程师 (25 题)

1. 在 Servlet 中可以通过设置___控制页面在客户端的缓存策略。

- A: response.setHeader("Pragma","No-cache");
- B: response.setDateHeader("Expires",0);
- C: response.setHeader("Cache-Control","no-cache");
- D: servlet.setCacheControl ()

AS

2. 下面哪个类用来执行存储过程的调用?

- A: CallableStatement
- B: PreparedStatement
- C: Statement
- D: 以上都是标准答案

3. 在 Java 程序中, 抛出下列哪个异常是需要特别声明?

- A: UnsupportedOperationException
- B: IllegalArgumentException
- C: OutOfMemoryError
- D: IOException

4. 关于 JSP 的描述正确的是:

- A: JSP 需要编译后才能执行
- B: JSP 是解释执行
- C: JSP 只能输出 HTML 页面
- D: JSP 是 JavaScript Page 的简称

5. 利用 Thread.wait() 同步线程, 可以设置超时时间吗?

- A: 能
- B: 不能

6. 以下代码的执行结果

```
static boolean foo(char c) {  
    System.out.print(c);  
    return true;  
}
```

```
public static void main(String[] args) {  
    int i;  
    for (foo('A'), i = 0; foo('B') && (i < 2); foo('C')) {  
        ++i;  
    }  
}
```



```

        foo('D');
    }
}
A: ABCDABCD
B: ABCDBCDB
C: ABDCBDCB
D: 运行时抛出异常
E: 编译错误

```

7.请问下列那种形式可用来注释一段 XML 内容?

- A: <? ?>
- B: <!-- -->
- C: <# #>
- D: <\$ \$>

8.现在流行的名词 IOC 的真实意思是:

- A: 依赖注入 (注入: Injection)
- B: 独立于容器 (独立: Independent, 容器: Container)
- C: 依赖倒置 (倒置、反转: Inversion)
- D: 控制倒置 (控制: Control, 倒置、反转: Inversion)

9.有关 final, finally, finalize 的描述, 正确的是

- A: final 用于声明成员变量, 方法和类, 分别表示成员变量不可变, 方法不可覆盖, 类不可继承等
- B: final 声明的成员变量, 必须在声明的同时赋值。
- C: try {} 里有一个 return 语句, 那么紧跟在这个 try 后的 finally{} 里的 code 在 return 后才会执行
- D: finalize 是 Object 类的一个方法, 在垃圾收集器执行的时候会调用被回收对象的此方法, 可以覆盖此方法提供垃圾收集时的相关资源释放

10.Open-Closed 原则的含义是指: 一个软件实体 ()

- A: 应当对扩展开放, 对修改关闭
- B: 应当对修改开放, 对扩展关闭
- C: 应当对继承开放, 对修改关闭
- D: 以上都不对

11.JUnit TestCase 基类中, 有一段代码:

```

public void runBare() throws Throwable {
    setUp();
    try {
        runTest();
    }
    finally {
        tearDown();
    }
}

```

```
}  
}
```

这段代码体现了哪一种设计模式？

- A: Template Methods
- B: Adapter
- C: Composite
- D: Command

12. 下面那些 method 不属于 javax.servlet.http.HttpServlet 类？

- A: doTrace
- B: service
- C: doUpdate
- D: doDelete

13. 设计 JDBC 连接池时需要考虑以下哪几个方面？

- A: 并发访问
- B: 连接池的分配与释放
- C: 数据库的性能
- D: 连接池的配置与维护

14. 关于继承表述错误的是：

- A: 继承是一种通过扩展一个已有对象的实现，从而获得新功能的复用方法。
- B: 泛化类（超类）可以显式地捕获那些公共的属性和方法。特殊类（子类）则通过附加属性和方法来进行实现的扩展。
- C: 继承会破坏封装性，因为会将父类的实现细节暴露给子类。
- D: 继承本质上是“白盒复用”，对父类的修改，不会影响到子类。

15. 对于 JDBC 连接，下面哪个表述是正确的？

- A: 连接是由 JDBC Driver 管理的
- B: 连接的建立及关闭极其耗资源
- C: 连接不需要在代码中显式关闭
- D: 连接是基于 RMI 创建的

16. 请问以下那类内容可以存放于 XML 文件中的 CDATA 区段？

- A: C 语言编译后的二进制内容
- B: 任意二进制内容
- C: 包含 <> 等特殊字符的文字内容
- D: 任意文字内容

17. wait()、notify() 不在 synchronized 方法或 synchronized block 中被调用的后果是：

- A: 没有问题，完全可以独立使用
- B: 编译通不过
- C: 可以编译，但运行时可能出错

18. 经过强制类型转换以后，变量 a, b 的值分别为 ()

short a = 128;

byte b = (byte) a;

A: 128 127

B: 128 -128

C: 128 128

D: 编译错误

19. XML Schema 中，对于 <xs:choice> 和 <xs:sequence> 的描述，那个是正确的？

A: 前者用于定义无顺序的元素，后者用于定义有顺序出现的元素

B: 前者定义可选择的元素，后者定义有序排列的元素

C: 前者定义选择元素，后者定义不可选择元素

D: 前者定义选择框，后者定义序列号

20. 领域模型是一组表示___，在设计工作中广泛用来启发设计软件对象。

A: 真实世界的概念类

B: 虚拟世界的概念类

C: 软件部件的模型

D: 硬件部件的模型

21. 有关 Servlet 的生命周期的说法，正确的是：

A: Servlet 被服务器实例化后，容器运行其 init 方法，请求到达时运行其 service 方法，service 方法自动派遣运行与请求对应的 doXXX 方法 (doGet, doPost)

B: 当服务器决定将实例销毁的时候调用其 destroy 方法。

C: servlet 处于服务器进程中，它通过多线程方式运行其 service 方法，一个实例可以服务于多个请求

D: Servlet 的实例在响应一次请求后会自动销毁

22. JDBC 的主要功能有？

A: A. 创建与数据库的连接

B: B. 发送 SQL 语句到数据库中

C: C. 处理数据并查询结果

D: D. 以上都是

23. 两个各有 n 个元素的有序列表并成一个有序表，其最少的比较次数是

A: n

B: 2n-1

C: 1

D: n-1

24.默认情况下, `connection.setAutoCommit()` 返回结果是?

- A: false
- B: true
- C: null
- D: 以上都不是

25.顺序执行下列程序语句后, b 的值是

```
String a = "Hello";  
String b = a.substring(1,2);
```

- A: el
- B: Hel
- C: He
- D: e

24.默认情况下, `connection.setAutoCommit()` 返回的结果是?

- A:false
- B:true
- C:null
- D:以上都不是

25.顺序执行下列程序语句后, b 的值是

```
String a = "Hello";  
String b = "a.substring(1,2);
```

- A:el
- B:Hel
- C:He
- D:e

B2B-技术部-工程师笔试试题（卷二）

卷二：搜索研发、c++（25题）

1. 关于 Document Object Model (DOM) 的描述哪些是正确的？

- A: DOM 可以描述 HTML 或者 XML
- B: Javascript 支持基于 DOM 的网页元素获取
- C: DOM 不支持对元素的修改
- D: 相对于基于 Tag 的网页处理，基于 DOM 的网页处理会占用较小的内存。

2. 网络蜘蛛系统主要的功能模块包括？

- A: 存储
- B: 调度
- C: 抓取
- D: 排重

3. 一个汉字在 UTF-8 编码中占用几个字节？

- A: 1
- B: 2
- C: 3
- D: 4

4. 一个网络蜘蛛系统需要避免重复抓取，下面哪些因素可作为确定重复性的依据？

- A: css
- B: 网页的风格
- C: url
- D: 网页内容

5. 已知 ξ 的分布列为

ξ -1 0 1

P 0.5 0.3 0.2

则 $E\xi$ 等于？

- A: 0
- B: 0.2
- C: -1
- D: -0.3

6. 小王在某个数据库中检索到了 50 篇文献，查准率和查全率分别为 40%、80%，则全部相关文档有___篇。

- A: 20
- B: 25
- C: 30
- D: 40

7. 下列关于索引压缩的说法哪些是正确的？
- A: 索引压缩减少了索引占用的空间
 - B: 索引压缩可以采用无损压缩编码或有损压缩编码
 - C: 在任何情况下索引压缩都会带来检索性能的提升
 - D: 索引压缩增加了查询时解压的时间消耗
8. 以下哪些属于设计 crawler 时提高抓取速度采用的策略？
- A: 多机并行抓取
 - B: 建立同步连接
 - C: 建立异步连接
 - D: 使用 DNS 缓存
9. 信息熵是描述事件信息量的度量，熵越大，则
- A: 事件越确定
 - B: 事件越不确定
10. Trie 树查询的时间复杂度最坏情况下是多少？
- A: $\log(n)$, n 是树的节点数
 - B: $\log(n)$, n 是查询词的长度
 - C: n , n 是查询词的长度
 - D: n , n 是树的节点数
11. 关于 HTML 的描述哪些是正确的？
- A: HTML 中指定网页编码的关键词是 charset
 - B: HTML 是一种自描述的语言
 - C: DIV 和 Table 都可以用来做页面布局
 - D: html 代码必须按照规范才能够在浏览器中显示
12. 关于 HTTP 协议的描述哪些是正确的？
- A: HTTP 协议一般是建立在 UDP 协议之上。
 - B: HTTP 协议不支持 persistent connection。
 - C: HTTP 是无状态的协议。
 - D: 服务器返回的 status code 在 200 到 300 之间，表示请求被成功处理。
13. 下列哪些属于信息检索模型？
- A: 布尔模型
 - B: 隐马尔科夫模型
 - C: 向量空间模型
 - D: 支持向量机模型

14. 下面哪些是常见的分布式通信协议？

- A: Java RMI
- B: CORBA
- C: ARP
- D: RPC

15. 下列哪些因素会影响一个分布式搜索引擎的效率？

- A: 服务器 cpu
- B: 服务器 io
- C: 网络状态
- D: 搜索算法优劣

16. 下面一段代码的输出是[]

```
void fn( int* b){
    (*b)++;
}
int main(){
    int a=7;
    fn(&a);
    cout<<a;
    return 0;
}
```

- A: 0
- B: 7
- C: 8
- D: undefined

17. 设森林 F 中有三棵树，第一，第二，第三棵树的结点个数分别为 M1, M2 和 M3。与森林 F 对应的二叉树根结点的右子树上的结点个数是[]

- A: M1
- B: M1+M2
- C: M3
- D: M2+M3

18. 有以下两个结构体，编译选项为 4 字节对齐，那么：

```
struct A{
    int a;
    short b;
    int c;
    char d;
};
```

```
struct B{
    int a;
```

```
short b;
char d;
int c;
};
```

候选答案:

- A: sizeof(A) = 16, sizeof(B) = 16
- B: sizeof(A) = 13, sizeof(B) = 12
- C: sizeof(A) = 16, sizeof(B) = 12
- D: sizeof(A) = 11, sizeof(B) = 11

```
19.int func(x)
{
    int count = 0;
    while(x)
    {
        count++;
        x = x&(x-1);
    }
    return count;
}
```

假定 x=9999, 函数返回值为:

- A: 6
- B: 7
- C: 8
- D: 9

20. 设双向循环链表的结点结构为:

```
struct node{
    void *data;
    struct node* lLink; //前一个节点指针
    struct node* rLink; //后一个节点指针
};
```

且不带表头结点, 若想在指针 p 所指结点之后插入指针 s 所指结点, 则应该执行下列那个操作 ()

- A: p->rLink=s; s->lLink=p; p->rLink->lLink=s; s->rLink=p->rLink;
- B: p->rLink=s; p->rLink->lLink=s; s->lLink=p; s->rLink=p->rLink;
- C: s->lLink=p; s->rLink=p->rLink; p->rLink=s; p->rLink->lLink=s;
- D: s->lLink=p; s->rLink=p->rLink; p->rLink->lLink=s; p->rLink=s;

21. 一组记录的关键码为 (46, 79, 56, 38, 40, 84), 则利用快速排序的方法, 以第一个记录为基准得到的一次划分结果为 ()

- A: (38, 40, 46, 56, 79, 84)
- B: (40, 38, 46, 79, 56, 84)
- C: (40, 38, 46, 56, 79, 84)
- D: (40, 38, 46, 84, 56, 79)

22. 关于 static, 下列说法错误的有:

- A: 在函数内部, static 局部变量只被初始化一次, 下一次依据上一次结果值
- B: 若全局变量仅由单个函数访问, 则可以将这个变量改为该函数的静态局部变量, 以降低模块间的耦合度
- C: 在源文件 A 中定义的 static 全局变量在源文件 B 中可以使用
- D: 静态全局变量是静态存储方式

23. 在 32 位系统下:

```
struct s1
{
    int i; 8;
    int j; 4;
    int a; 3;
    double b;
};
struct s2
{
    int i; 8;
    int j; 4;
    double b;
    int a; 3;
};
```

sizeof(s1) 和 sizeof(s2) 为 ()

- A: 20, 20
- B: 16, 24
- C: 16, 20
- D: 12, 16

24. vector<iterator> 重载了下面哪些运算符?

- A: ++
- B: >>
- C: * (前置)
- D: ==

25. 若有以下定义和语句: int w[2][3], (*pw)[3]; pw=w; 则对 w 数组元素的非法引用是()

- A: *(pw+1)[2]
- B: *(w[0]+2)
- C: pw[0][0]
- D: *(pw[1]+2)

B2B-技术部-工程师笔试试题（卷三）

卷三：公共题（17题）

1. 如果一个判定中的复合条件表达式为 $(A > 1) \text{ OR } (B \leq 3)$ ，则为了达到 100% 的覆盖率，至少需要设计多少个测试用例

- A: 1
- B: 2
- C: 3
- D: 4

A 2. 典型的针对系统漏洞的 DoS 攻击方式是哪种？

- A: Syn flood *Denial of Service 拒绝服务*
- B: Smurf
- C: Ping of Death
- D: UDP flood

D 3. 根据前面的数字，括号中应该是：4，2，2，3，6，15，（ ）

- A: 20
- B: 24
- C: 25
- D: 45

4. 有三个桶，容量分别是 3 升，5 升，7 升，你只能进行下面的操作：把一个桶中所有的水倒掉；把一个桶 A 中的水倒入桶 B，直到桶 A 空了或者桶 B 满了；假设一开始容量为 3 升和 5 升的桶是满的，7 升的桶是空的，希望通过一系列操作使 3 个桶中任意一个中正好有 4 升水，那么最少需要 [] 次操作？

- A: 2
- B: 3
- C: 4
- D: 5

5. 在一次实验中，研究人员将大脑分成若干个区域，然后扫描并比较了每个人大脑各区域的脑灰质含量。结果显示，智商测试中得分高的人与得分低的人相比，其大脑中有 24 个区域灰质含量更多，这些区域大都负责人的记忆、反应和语言等各种功能。从这段文字中，我们可以推出

- A: 智商低的人大脑中不含灰质
- B: 大脑中灰质越多的人，智商越高
- C: 聪明的人在大脑 24 个区域中含有灰质
- D: 智商高的人，记忆、反应和语言能力都强

6.某报评论：H市的空气质量本来应该已经得到改善。五年来，市政府在环境保护方面花了气力，包括耗资600多亿元将一些污染最严重的工厂迁走，但是，H市仍难摆脱空气污染的困扰，因为解决空气污染问题面临着许多不利条件，其中，一个是机动车辆的增加，另一个是全球石油价格的上升。以下各项如果为真，都能削弱上述论断，除了

- A：近年来H市加强了对废气的排放的限制，加大了对污染治理费征收的力度。
- B：近年来H市启用了大量电车和使用燃气的公交车，地铁的运行路线也有明显增加。
- C：近年来H市启用了大量电车和使用燃气的公交车，地铁的运行路线也有明显增加。
- D：由于石油涨价，在国际市场上一些价位偏低的劣质含硫石油进入H市
- E：由于汽油涨价和公车改革，拥有汽车的人缩减了驾车旅游的计划。

7.用户进程的跟踪信息存在哪个目录？

- A: cdump
- B: bdump
- C: udump
- D: vdump

8.下列软件中，哪种可以保护邮件传输安全？

- A: Gnupg
- B: TureCrypt
- C: WinwebMail
- D: SendMail

9.使普通用户可执行超级用户文件的命令是

- A: chmod +v
- B: usermod +v
- C: chmod +s
- D: usermod +s

10.软件测试的对象包括()

- A: 目标程序和相关文档
- B: 源程序、目标程序、数据及相关文档
- C: 目标程序、操作系统和平台软件
- D: 源程序和目标程序

11.下列关于软件缺陷生命周期的状态定义的说法中，错误的是()

- A：“新建”表示测试中发现并报告了新的软件缺陷
- B：“打开”表示软件缺陷已被确认并分配给相关开发工程师处理
- C：“关闭”表示软件缺陷已被开发人员修复并等待测试人员验证
- D：“解决”表示软件缺陷已被确认修复

12. 随着 Google 逐步开放其互联网应用程序的 API, 包括 Amazon、Facebook、MySpace、阿里巴巴集团等一批著名的互联网企业将自己的应用接口开放出来, 供第三方开发商定制服务。为了确保 OpenAPI 平台质量, 下列不是平台提供商 (如阿里巴巴) 的测试内容

- A: 安全测试
- B: 单元测试
- C: 验收测试
- D: 性能测试

13. 向有限的空间输入超长的字符串是下列攻击中的哪一种?

- A: 网络监听
- B: 端口扫描
- C: 缓冲区溢出
- D: IP 欺骗

14. 项目最重要的特性是 () 和 () ?

- A: 独特性
- B: 可管理性
- C: 一次性
- D: 可重复性
- E: 可预测性

15. Web 应用安全漏洞最突出的有: Cross Site Scripting, SQL injection 等。Cross Site Scripting 是一种存在 Web 应用中, 允许黑客向 Web 应用注入恶意脚本, 以获取其他用户重要数据和隐私信息为目的的一种攻击形式。下面哪个输入可能导致 Cross Site Scripting

- A: `SELECT count(*) FROM users_list_table WHERE username='USERNAME' AND password='PASSWORD' or 1=1`
- B: `<script>alert(document.cookie);</script>`
- C: `
crack you!</br>`
- D: `http://www.alibaba.com.cn/index.php?dir=../../usr/local/apache2/conf/`

16. 简述隐马尔可夫模型 (HMM) 的三个基本问题

评估 解码 学习

17. 有两个有序整数集合 a, b. 请写一个函数, 实现找出 a, b 集合中的交集, 并打印出来