

关于Java编程的中文问题的几条分析原则[1] PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/244/2021_2022__E5_85_B3_E4_BA_8EJava_c104_244672.htm 尽管关于Java中文问题的讨论已经相当多了，但由于Java的相关技术标准繁多，面向Java的Web服务器、应用服务器以及JDBC数据库驱动等都没有官方的标准，所以Java应用在处理中文时所存在的问题不仅没有消失而且随着所选用的服务器、驱动程序以及运行环境等因素的不同而变化。那么我们如何从众多现象中找出问题所在，并进行分析和解决呢？与大部分的讨论不同，本文将主要从如何预测、发现和检查问题的角度给出建议，帮助开发人员找出可能引起问题的各种源头，从而更好地解决Java的中文问题。

引言 尽管对于Java中文处理问题的讨论已不乏其数，但由于Java技术涉及内容广（J2EE包含了十几种相关技术），技术供应商繁多，面向Java的Web服务器、应用服务器以及JDBC数据库驱动等都没有官方的标准，所以Java应用在处理中文过程中出了存在固有的问题外也存在随着选用的服务器，驱动程序的不同而带来的Java中文问题的多变性，增加了问题的复杂度。那么，我们如何在这么纷繁的现象中找到问题的症结呢？Java中文问题的一般解决办法事实上，Java的中文问题都是由于Java应用所采用的缺省编码格式与目标或者应用所要读入字符的编码格式不同而造成的（具体参见文献1）。

对于如何解决Java的中文问题，通常有四种方法：1）选择JDK的中文本地化版本。尽管Java2 JDK的中文本地化版本（<http://java.sun.com/products/jdk/1.2/chinesejdk.html>）并不是一个官方的版本，Sun公司也没有承诺会对该本地化版本进行

升级，但其仍不失为一个Java中文问题的解决方案。2) 选择合适的编译参数。对于Java的国际版本来讲，我们也可以在编译Java应用的时候通过指定确定的编码机制来实现其编译结果对中文的支持。例如，对于需要支持繁体中文和简体中文应用可以通过javac -encoding big5 sourcefile.java 和javac -encoding gb2312 sourcefile.java来编译源程序。3) 通过编程的方式实现字符编码的转换代码。通过编程的方式来解决Java的中文问题，已经成为了一种较为普遍的做法。下面就是一种最常见的字符编码转换函数，其将字符的编码格式转换为中文Windows系统的GBK编码形式。

```
public static String toChinese(String strvalue){try{if(strvalue==null)return null.else{strvalue = new String(strvalue.getBytes("ISO8859_1"), "GBK").return strvalue.}}catch(Exception e){return null.}}
```

4) 定义字符输出集。对于JSP应用，我们可以通过或来定义JSP页面的字符输出集。当然，我们也可以通过HTML的标记来定义字符的输出集。存在的问题 根据方法实现的方式，我们可以将以上四种方法分为两类，一类是通过利用某些标准或者规则来实现的方法，上面的1)、2)、4) 都属于此类；一类是通过针对性的编程来实现的方法，上面所提的方法3) 就属于此类。由于方法1), 2), 4) 是具有规范性的一类方法，所以方法比较简单，解决方案也不具备较大的针对性，较为通用，例如我们可以采用方法2) 的编译方式通过编译Java源文件来实现内码的预置，而无需考虑源码到底有哪些部分出现了Java的中文处理问题，诸如输出乱码等等。但是，正由于这些方法不具备针对性，解决问题的方法过于统一，所以在某些情况下，它们并不能彻底地解决Java的中文问题。举一个非常常见

的例子。在通常情况下，用户的Java应用往往需要与其它Java应用接口进行交互，例如通过某种版本的JDBC访问数据库。由于JDBC的驱动所支持的编码随着提供商乃至版本的不同而不同，所以如果在数据库的输入输出过程中出现中文不能正确处理问题时，我们需要在数据的输入和输出过程做两次正好相反的编码转换，这对于方法1)，2)，4)来说，往往是无法解决的。当然，对于方法2，我们也可以通过采用一些技巧使来满足上面的情况，一个最有效的办法就是尽量将Java应用的各个部分组件化。例如我们可以通过将数据库的读入和输出代码分解在不同的源文件上来实现分别编译，从而满足不同的字符编码要求。但是通常的程序设计都不太可能满足这种要求，因为这种程序的划分结果很可能是不合理的。例如，我们将数据库的读出和写入方法封装到一个类中是比较合适的一种设计，但如果将该类的这两个方法分别实现在两个文件里则变得非常不合理。因此对于1)，2)，4)方法来说，虽然实现比较简单，但却具有一些无法克服的缺点。这也是那些实现起来相对复杂的编程方法得以流行的原因。

100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com