

在JAVA开发中的中文处理问题及解决办法Java认证考试 PDF
转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/620/2021_2022__E5_9C_A8J_AVA_E5_BC_80_c104_620971.htm 在应用开发的过程中，总会遇到一些“难以理解”的系统缺陷和“不易解决”的问题。其实，通过认真分析，不必购买昂贵的产品依然可以解决多数问题。 JDBC ODBC Bridge的Bug及其解决方法 在编写一数据库管理程序时，发现JDBC-ODBC Bridge存在不易发现的Bug。在向数据表插入数据时，如果为英文字符，存储内容完全正确，如果存入中文字符，部分数据库只能存储前七八个中文字符，其他内容被截去，导致存储内容的不完整（有些数据库不存在这个问题，如Sybase SQL Anywhere 5.0。 JDBC-ODBC Bridge还存在无法建表的Bug）。对于广大需要存储中文信息的Java程序员来说，这可是一个不好的消息。要么改用其他语言编程，要么选择其他价格昂贵的数据库产品。“一次编写，到处运行”的目标，也大打折扣。能不能采用变通的方法，将中文信息进行处理后再存储来解决这个问题呢？答案是肯定的。解决问题的具体思路、方法 Java采用Unicode码编码方式，中英文字符均采用16bit存储。既然存储英文信息是正确的，根据一定规则，将中文信息转换成英文信息后存储，自然不会出现截尾现象。读取信息时再进行逆向操作，将英文信息还原成中文信息即可。由GB2312编码规则可知，汉字一般为二个高位为1的ASCII码，在转换时将一个汉字的二个高位1去掉，还原时再将二个高位1加上。为了处理含有英文字符的中文字串，对英文字符则需要加上一个Byte 0标记。以下提供的两个公用静态方法，可加入任何一个

个类中使用。 将中英文字符串转换成纯英文字符串

```
public static String toTrueAsciiStr(String str){ StringBuffer sb = new StringBuffer(). byte[] bt = str.getBytes(). for(int i=0 .i  bt.length.i ){ if(bt[i]  0){ //是汉字去高位1 sb.append((char)(bt[i]&0x7f)). }else{//是英文字符 补0作记录 sb.append((char)0). sb.append((char)bt[i]). } } return sb.toString(). }
```

将经转换的字串还原

```
public static String unToTrueAsciiStr(String str){ byte[] bt = str.getBytes(). int i,l=0,length = bt.length,j=0. for(i = 0.i  length.i ){ if(bt[i] == 0){
```

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