

让你不苦恼：Java的中文编程配置心得 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/275/2021_2022__E8_AE_A9_E4_BD_A0_E4_B8_8D_E8_c67_275163.htm Java的中文编程与配置心得 Java的中文问题历史悠久，连绵不绝，至今也没有完全解决，但是上有政策下有对策，我们总是有办法搞定它的。跟Java相关的中文问题主要有两类，一类是编程的问题，涉及到I/O，内码转换等。第二类是Java运行环境的配置，涉及字体，属性配置等。我刚刚用了一天的时间解决这些问题，觉得很有必要给自己写个备忘录之类的。我看还是从问题入手吧，这样不致于让大家打瞌睡。我想写个程序，这个程序有个基本功能就是显示文件内容，我用JTextArea来做显示的事情，程序简单的到家了，但是就是中文都是乱码。我的配置是JBuilder7，JBuilder自带的JDK1.3.3_b24，我自己装的JDK是JDK1.4.0_02_b02，都是主流的JDK。操作系统是英文Windows2000加中文支持包。我尝试换JDK，1.3.3和1.4.0都不行，down一个最新的j2sdk-1_4_1-rc也是不行，好像不是JDK的问题，于是我就把精力集中到I/O的编码转换上，我查阅了网上若干关于JAVA中文问题的文章，把编码转换搞的倍儿清楚，可是怎么试，换什么编码折腾都不行，反而显示的更糟糕，当初还是乱的有些规矩，就是象在纯英文系统上显示的那样，好歹我还知道那是汉字，只是显示不出来，等我折腾编码，就变得都是问号了。编码转换心得：JAVA内部是UNICODE编码，在I/O时如果使用Reader/Writer就要发生编码转换，使用系统属性file.encoding作为编码方式。如果使用Stream就没有转换的事情了，那是Binary的数据。有用的

方法有：1。在Reader/Writer上加encoding的选项，注意编码的方向，在Reader中的encoding表示把数据从encoding转换成Unicode，writer就是把Unicode的字符转换成encoding格式的。2。用String.getBytes()把字符串转换成指定编码。常用的编码格式：ISO8859_1，这个是英文系统缺省的8bit编码，因为是8bit的，所以不会把汉字的高位删去，所以用它也是可以处理汉字的(我自己这么理解，总觉得有些不妥，但又不知道不妥在何处，还望高人指出)。GB2312和GBK，汉字编码，推荐使用GBK，它兼容GB2312并且支持更多汉字。UNICODE，一个大的字符集，不知是不是国际标准反正大家都支持，使用16位对每个字符编码，汉字虽然正合适，但英文却吃亏了，要用多一倍的空间来存储，所以很多人还是老大不乐意，写的程序不支持UNICODE。JSP/Servlet的中文问题有两种解决办法：1。不在程序中进行编码转换，把这个工作交给浏览器，方法就是用javac encoding GBK *.java来编译所有的bean，然后在JSP页面上加 Word-WRAP: break-Word" bgColor=#f3f3f3>以下是引用片段：或者是在HTML中直接加：到底加那个，试试就知道了，我也搞不清楚了。2。在程序中指定编码，用javac encoding ISO8859_1 *.java来编译所有的bean，在涉及到中文显示的程序上加 str=new String(str.getBytes("ISO8859_1")).上面两种方法不能混用，意思就是要么就是GBK，要么就是ISO8859_1，从里到外都一样就好了。数据库JDBC的中文问题，一般只要按照数据库指定的编码进行转换，比如按照ISO8859_1读，ISO8859_1写，一般就没什么问题了。虽然有这些编码上的心得，但是并不能解决我的问题。看来我的程序输入输出用的都是ISO8859_1，

我的问题跟编码没什么关系。是不是字体的问题呢?在Swing的组件中，字体总是那么几个，基本上是定死的，选那个都不行。但是我突然发现可以更改这些字体的配置，就是font.properties这个文件，一般JDK都带了中文的字体配置文件，可能是font.properties.zh之类的，不同版本的JDK名字有些差别，你要做的就是用中文的配置覆盖font.properties文件。我满心欢喜的以为成功了，但是失败无情的又一次打击了我。不是这种方法不对，但是在Windows系统中，java能够比较自动的检查你的系统编码，使用最合适的字体配置文件，一般不需要你改动了，在JDK1.2之前确实是要这么改的，难怪那篇文章是JDK1.1的文档呢。连font.properties也不灵了，唉，是不是JBuilder有问题呀?到了这儿，我不得不怀疑它了，虽然它从没让我失望过，我关掉JBuilder，拿起JDK命令行，哎呀!可爱的中文竟然出来了!简直是.....莫名其妙呀，因为JBuilder也是用的这个JDK呀?没理由呀，我再次打开JBuilder，运行，乱码!怪事.....我来看看.....我把JBuilder里面的执行命令拷贝到DOS窗口里手动执行....没有中文.....怪怪怪，我仔细对照我手敲的命令和JBuilder拷出来的命令，除了我用的是java它用的是javaw之外，别的几乎都一样，难道.....这个java和javaw竟然会不一样?唉.....要我说什么呢，答案就是这个了，javaw会使用与java不同的Local配置进行工作，导致了我一直看不到中文，而java就是好好的。JBuilder也是好好的，没有问题，调用javaw也不是它的错，难道你想每次运行或者调试的时候蹦出个DOS黑框么?呵呵。这是JDK的一个Bug，我在java的网站上查到这个bug，描述的情况跟我的一样，然后Sun说这个Bug已经修复了，嘿嘿，简直就是扯淡!修复个

茄子!Bug编号是：4629351。不信可以去看看。但是难道我就不能在JBuilder里面看我可爱的中文了吗?好像JBuilder5的时候还能选择是用java还是javaw，现在好像没法儿选了。怎么才能让javaw支持中文呢?经过又是一番苦苦查找，竟然没有什么结果，郁闷!正当苦闷之时，突然想起当初配置过Unix下的TOMCAT，好像加了些启动参数，就能让tomcat支持中文了，赶快翻出以前的文档，啊，幸好我当初写了工作心得，要不然又抓瞎了。其实道理很简单，就是给JAVA虚拟机加上一些属性：以下是引用片段：-Dfile.encoding=GBK
-Ddefault.client.encoding=GBK -Duser.language=zh
-Duser.region=CN 把它跟在javaw的命令行后面当参数!
在JBuilder中选择Project->Project Properties->run->edit->VM Parameters，填上就行了!配置心得：在JDK低版本和Unix下，需要改动font.properties，让JAVA VM能够找到合适的字体来显示汉字。目前为止，javaw有Bug，只使用英文的Local，需要给VM加上属性才行正常显示中文。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