

从GB2312到Unicode转换表制作程式 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/221/2021_2022__E4_BB_8EG B2312_E5_c104_221449.htm 简体中文有两种常用编码

: GB2312 和 Unicode. 它们的对应关系可以用下面的程式所生成:

```
/** * GB2312Unicode.java * Copyright (c) 2003 by Dr. Herong Yang, http://www.herongyang.com/ */ import java.io.*. import java.nio.*. import java.nio.charset.*. class GB2312Unicode { static OutputStream out = null. static char hexDigit[] = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F}. static int b_out[] = {201,267,279,293,484,587,625,657,734,782,827, 874,901,980,5590}. static int e_out[] = {216,268,280,294,494,594,632,694,748,794,836, 894,903,994,5594}. public static void main(String[] args) { try { out = new FileOutputStream("gb2312.gb"). writeCode(). out.close(). } catch (IOException e) { System.out.println(e.toString()). } } public static void writeCode() throws IOException { boolean reserved = false. String name = null. // GB2312 is not supported by JDK. So I am using GBK. CharsetDecoder gbdc = Charset.forName("GBK").newDecoder(). CharsetEncoder uxec = Charset.forName("UTF-16BE").newEncoder(). CharsetEncoder u8ec = Charset.forName("UTF-8").newEncoder(). ByteBuffer gbbb = null. ByteBuffer uxbb = null. ByteBuffer u8bb = null. CharBuffer cb = null. int count = 0. for (int i=1. i=1 amp. i=10 amp. i=16 amp. i=56 amp. i=88 amp. i"). writeNumber(i). writeString(" Row: " name). writeln(). writeString(""). writeln(). if (!reserved) { writeln(). writeHeader(). // looping through all characters in one row for (int
```

```
j=1. j> 4) amp. 0x0F]). } public static void writeHeader() throws  
IOException { writeString(""). writeln(). writeString("Q.W. ").  
writeGBSpace(). writeString(" GB Uni. UTF-8 "). writeString(" ").  
writeString("Q.W. "). writeGBSpace(). writeString(" GB Uni. UTF-8  
"). writeln(). writeln(). } public static void writeFooter() throws  
IOException { writeString(""). writeln(). } public static boolean  
validGB(int i,int j) { for (int l=0. l=b_out[l] amp. i*100 j程式输出的
```

列表格式如下： Q.W. GB Uni. UTF-8 Q.W. GB Uni. UTF-8
1601 啊 B0A1 554A E5958A 1602 阿 B0A2 963F E998BF 1603 埃
B0A3 57C3 E59F83 1604 挨 B0A4 6328 E68CA8 1605 哎 B0A5
54CE E5938E 1606 唉 B0A6 5509 E59489 1607 哀 B0A7 54C0
E59380 1608 皑 B0A8 7691 E79A91 1609 癌 B0A9 764C E7998C
1610 蔼 B0AA 853C E894BC 1611 矮 B0AB 77EE E79FAE 1612 艾
B0AC 827E E889BE 1613 碍 B0AD 788D E7A28D 1614 爰 B0AE
7231 E788B1 1615 隘 B0AF 9698 E99A98 1616 鞍 B0B0 978D
E99E8D 1617 氨 B0B1 6C28 E6B0A8 1618 安 B0B2 5B89 E5AE89
1619 俺 B0B3 4FFA E4BFBA 1620 按 B0B4 6309 E68C89 1621 暗
B0B5 6697 E69A97 1622 岸 B0B6 5CB8 E5B2B8 1623 胺 B0B7 80FA
E883BA 1624 案 B0B8 6848 E6A188 1625 肮 B0B9 80AE E882AE
1626 昂 B0BA 6602 E69882 1627 盎 B0BB 76CE E79B8E 1628 凹
B0BC 51F9 E587B9 1629 敖 B0BD 6556 E69596 1630 熬 B0BE 71AC
E786AC 1631 翱 B0BF 7FF1 E7BFB1 1632 袄 B0C0 8884 E8A284
1633 傲 B0C1 50B2 E582B2 1634 奥 B0C2 5965 E5A5A5 1635 懊
B0C3 61CA E6878A 1636 澳 B0C4 6FB3 E6BEB3 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