

# 网页编码之GB2312、GBK与UTF-8的区别

原创

陆小马

于 2021-11-10 20:10:12 发布

1509

收藏 2

版权

分类专栏:

学习笔记

文章标签:

字符编码



学习笔记 专栏收录该内容

2 订阅

59 篇文章

订阅专栏

首先，我们要明白，GB2312、GBK和UTF-8都是一种字符编码，除此之外，还有好多字符编码。只是对于我们中国人的网站来说，用这三种编码 比较多。简单的说一下，为什么要用编码，在计算机内，储存文本信息用ASCII码，每一个字符对应着唯一的ASCII码。最初计算机是由美国发明的，他们也用的是键盘和上面的字母，所以他们的字符ASCII好解决。但是我们中国 的就不同了，每个汉字要对应唯一的ASCII码。这样，就出来了国家制定的字符编码标准：GB2312、GBK等。其他国家，其他语言也有他们对应的编码 标准。

GB 就是国标的意思，GB2312和GBK主要用于汉字的编码，而UTF-8是全世界通用的。意思就是说，如果你的网页主要面对使用汉语的中国人的话，使用 GB2312和GBK非常好，文字储存体积小，有一些优点。如果你的网页要面向世界的话，你再用GB2312和GBK作为网页编码的话，有些电脑上的浏览器没有这种编码，你的网页汉字内容就会变成无法识别的乱码。它们通常用在网页的meta标签内，例如：，表示这个页面使用的是GB2312编码。这个信息是给浏览器看的，浏览器会优先考虑使用从网页头部提取出来的编码信息对网页进行解码。当然，我们也可以强制浏览器使用某种编码解释网页，这样我们就看到了传说中的乱码。

GBK、GB2312等与UTF8之间都必须通过Unicode编码才能相互转换：

GBK、GB2312——Unicode——UTF8

UTF8——Unicode——GBK、GB2312

对于一个网站、论坛来说，如果英文字符较多，则建议使用UTF-8节省空间。不过现在很多论坛的插件一般只支持GBK。

如果是中文的网站 推荐GB2312 GBK有时还是有点问题 为了避免所有乱码问题，应该采用UTF-8，将来要支持国际化也非常方便 UTF-8可以看作是大字符集，它包含了大部分文字的编码。

使用UTF-8的一个好处是其他地区的用户（如香港台湾）无需安装简体中文支持就能正常观看你的文字\*而不会出现乱码。\*

gb2312是简体中文的码  
gbk支持简体中文及繁体中文  
big5支持繁体中文  
utf-8支持几乎所有字符

中国大陆最常用的就是GBK18030编码，除此之外还有GBK，GB2312，这几个编码的关系是这样的。 最早制定的汉字编码是GB2312，包括6763个汉字和682个其它符号 95年重新修订了编码，命名GBK1.0，共收录了21886个符号。 之后又推出了GBK18030编码，共收录了27484个汉字，同时还收录了藏文、蒙文、维吾尔文等主要的少数民族文字，现在WINDOWS平台必需要支持GBK18030编码。

GB2312编码大约包含6000多汉字（不包括特殊字符），编码范围为第一位b0-f7,第二位编码范围为a1-fe(第一位为cf时,第二位为a1-d3),计算一下汉字个数为6762个汉字。当然还有其他的字符。包括控制键和其他字符大约7573个字符编码 gbk编码是对GB2312编码的扩充，容纳的汉字更多，但仅仅是扩充，没有质的变化。保留了所有GB2312编码,在此基础上进行编码范围的扩充.容纳(包含特殊字符)共22014个字符编码. gb18030编码是在gbk编码基础上的扩充，因为汉字更多，仅仅使用两位编码已经不能容纳要求的汉字，所以采用了2+4位混和的办法，可以支持更多的汉字编码。并且保留了原有的gbk 2字节编码兼容GB2312和gbk编码的文件。大概容纳55657个编码(包含特殊字符) unicode编码(也就是UTF编码)：俗称万国码，致力于使用统一的编码准则表达各国的文字。为表达更多的文字，utf-8采用2/3混编的方式。目前容纳的汉字范围小于gbk编码。并且以3字节的方式处理中文，带来了兼容性的问题，原有的gbk,GB2312,gb18030编码文件都不能正常的处理,还有很长的路要走。

### gbk和gb2312的区别有哪些

首先大家要了解什么是gbk?什么是gb2312? 我们要知道他们都是一种字符编码，当然字符编码还有很多种。

而字符编码我们可以这样理解：

在计算机中存放的都是0和1的二进制值。

8个位对应一个字节，常用16进制来表示。

那么我们如果想要在计算机上看到我们想要的字符显示，而不是各种0和1的数字该怎么实现呢？

这里我们就需要使计算机把其所存储的对应的16进制的数值，转化为对应的字符，包括英文和中文等其他语言的字符，然后输出到屏幕上。

所以编码也就是，定义了一套规则，去指定哪些数值，对应着哪些字符。

那么字符编码，就是定义了一套规则，指定了计算机中存放的这么多值中的哪个值，对应了电脑屏幕显示出来的哪个字母。

综上所述，大家应该都能理解GBK和GB2312是一种字符编码了吧。

下面我们再具体说说他们的区别与相同点：

相同点：

- 1、GBK和GB2312 都是16位的！
- 2、它们通常用在网页的meta标签内。

不同点：

- 1、GBK字符编码支持简体中文和繁体中文！

GBK全称《汉字内码扩展规范》（GBK即“国标”、“扩展”汉语拼音的第一个字母，英文名称：Chinese Internal Code Specification），中华人民共和国全国信息技术标准化技术委员会1995年12月1日制订，国家技术监督局标准化司、电子工业部科技与质量监督司1995年12月15日联合以技监标函1995 229号文件的形式，将它确定为技术规范指导性文件。

## 2、GB2312只支持简体中文！

《信息交换用汉字编码字符集》是由中国国家标准总局1980年发布，1981年5月1日开始实施的一套国家标准，标准号是GB 2312—1980。

GB 2312标准共收录6763个汉字，其中一级汉字3755个，二级汉字3008个；同时，GB 2312收录了包括拉丁字母、希腊字母、日文平假名及片假名字母、俄语西里尔字母在内的682个全角字符。

如果你的网页主要面对使用汉语的中国人的话，使用 GB2312和GBK非常好，文字储存体积要小，有一些优点。如果你的网页要面向世界的话，你再用GB2312和GBK作为网页编码的话，有些电脑上的浏览器没有这种编码，你的网页汉字内容就会变成无法识别的乱码。

网页编码为什么不用gbk或者gb2312而用utf-8 09-22

GBK的文字编码是双字节来表示的，至于UTF-8编码则是用以解决国际上字符的一种多字节编...

gbk utf8如何选择 正确理解和使用GBK及UTF-8网页编码 12-13

网页编码英文译为web page encoding，是在网页中指定其特定的字符编码格式的库。GBK是...

UTF-8 GBK GB2312 之间的区别和关系\_utf8 gbk gb2312\_Robben.Han的博... 7-31

UTF8-Unicode-GBK、GB2312 DIVCSS5简单从功能上说: 1、GBK通常指GB2312编码 只支持...

字符编码GBK、GB2312和UTF-8的区别与联系\_菜鸟之小菜的博客 8-9

GBK、GB2312以及UTF-8就是在这种条件下诞生的。先来看看GB2312和GBK这两种中文编码...

在 HTML 中声明 UTF-8 编码方式 科大云炬 637

属性指定了编码方式为 UTF-8。这样浏览器就会按照 UTF-8 的方式解析网页内容，确保网页中...

GBK、UTF-8编码转换工具 11-01

GBK、UTF-8批量文件3秒快速转码工具(支持GBK,UTF-8免费转换),UTF-8/GBK编码在线转换工...

常见编码格式UTF-8 GBK GB2312 ASCII Unicode 的区别\_一棵栗子树的博客... 8-3

UTF-8是Unicode的实现方式之一,它对不同范围的字符使用不同长度的编码。1)对于单字节的...

...8编码介绍\_unicode utf8 gbk gb2312\_0溺水的鱼0的博客 8-2

然后,从“严”的最后一个二进制位开始,依次从后向前填入格式中的x,多出的位补0。这样就得到...

把页面搜索URL中的UTF-8编码还原为汉字 无知的我 2607

在进行页面浏览的时候，能够看到地址栏中的URL很少有显示中文字符的，全是符号和英文字...

UTF-8 GBK UTF8 GB2312 之间的区别和联系 java开发者的专栏 855

UTF-8: Unicode TransformationFormat-8bit, 允许含BOM, 但通常不含BOM。是用以解决国...

UTF-8、GBK、GB2312、GB18030、ASCII编码基本关系\_gb2312和gb18030... 8-11

兼容性关系是GB18030兼容GBK,GBK兼容GB2312,GB2312兼容ASCII。所谓兼容,你可以简单理...

GB2312、GBK与UTF-8的区别\_阳光\_你好的博客 8-7

GB 就是国标的意义,GB2312和GBK主要用于汉字的编码,而UTF-8是全世界通用的。意思就是说,...

GB2312、GBK与UTF-8的区别 赵圆圆 3万+

首先，我们要明白，GB2312、GBK和UTF-8都是一种字符编码，除此之外，还有好多字符编码...

utf-8、gbk、gb2312的关系 lf\_1234的博客 561

一句话来说：urf-8(世界文字) > gbk(中国汉字) > gb2312(gbk的子集) UTF-8包含全世界所有国家...

系统运维系列 之编程中UTF-8、GBK和GB2312等编码格式之间的区别和联系... 8-2  
1.3GBK 编码: GBK是GB2312的扩展,是向上兼容的,因此GB2312中的汉字的编码与GBK中汉字的...

编码方式gbk、gb2312、utf8的区别 感恩、奋进、自信 860  
gbk一般用于繁体中文,是国家标准gb2312基础上扩容后兼容gb2312的标准。文字编码(中英...

网络编码 GB2312、GBK与UTF-8的区别 aizhejuan4396的博客 450  
网络编码 GB2312、GBK与UTF-8的区别 GB2312、GBK与UTF-8的区别这是一个异常经典的问...

字符编码: GBK、GB2312、UTF-8 686  
一、各编码间的区别 UTF-8: Unicode TransformationFormat-8bit, 允许含BOM, 但通常不含...

GB2312转UTF-8 (C代码) 12-04  
本资源为GB2312转UTF-8的C语言代码,实现方式为查表的方式,提供了可以调用的接口,里...

GB2312编码和UTF-8互转 (c语言实现) 01-01  
GB2312编码与utf-8编码的字符串的转换,主要使用windows api函数MultiByteToWideChar和...

GB2312, GBK, UTF-8的关系 pureboy的专栏 555  
1, GBK编码兼容GB2312.详细点的信息是: GB码, 全称是GB2312-80, GB码共收录6763个简...

让浏览器以utf-8编码格式解析text/plain的文件 huyao的博客 6486  
只需再响应头中设置编码格式即可 res.setHeader('Content-Type', 'text/plain; charset=utf-8'); ...

GBK编码的理解 sgmcy的专栏 1922  
GBK编码的理解

GBK 编码 热门推荐 极客神殿 5万+  
GBK编码范围: 8140—FEFE, 汉字编码范围见第二节: 码位分配及顺序。GBK编码, 是对GB2...

字符编码GBK、GB2312和UTF-8的区别与联系 weixin\_34092370的博客 106  
在计算机系统内, 文本信息的存储使用的ASCII码。什么是ASCII码呢? ASCII码全称是“美国信...

java utf-8 gbk\_Java 字符转码之UTF-8转为GBK/GB2312 最新发布 06-02  
在 Java 中, 可以使用`String`类的`getBytes()`方法将字符串转换为指定编码的字节数组, 然...

“相关推荐”对你有帮助么?

 非常没帮助  没帮助  一般  有帮助  非常有帮助

关于我们 招贤纳士 商务合作 寻求报道 400-660-0108 kefu@csdn.net 在线客服 工作时间 8:30-22:00  
公安备案号11010502030143 京ICP备19004658号 京网文〔2020〕1039-165号 经营性网站备案信息  
北京互联网违法和不良信息举报中心 家长监护 网络110报警服务 中国互联网举报中心 Chrome商店下载  
账号管理规范 版权与免责声明 版权申诉 出版物许可证 营业执照 ©1999-2023北京创新乐知网络技术有限公司



陆小马  
码龄8年 暂无认证

264 3万+ 3433 85万+  
原创 周排名 总排名 访问 等级

6706 581 295 73 1544  
积分 粉丝 获赞 评论 收藏



私信

关注

搜博主文章



## 热门文章

创业者必备知识SWOT分析模型+案例分析

34049

无线网络技术有哪些-智能家居，物联网

31308

物联网（IoT）及其未来应用方向

27963

Flash中传统补间，补间动画，补间形状的区别

23229

单片机蜂鸣器源代码+仿真

19384

## 分类专栏

	技巧经验软件网站...	付费	54篇
	农业工程与信息技术	付费	36篇
	黑客网络安全	付费	28篇
	Java,springboot技...	付费	40篇
	学习笔记		59篇
	Python		12篇



## 最新评论

单片机按键识别c语言程序+仿真

Y33787: 仿真图没有了么？

2. 基于顺序存储结构的图书信息表的排序

南拉河胖橘: 为啥我运行出来不对呢？

单片机蜂鸣器源代码+仿真

(^o^)-☆°.\*.°: 程序全部看懂了，但是我移植到我的程序仿真里还是不...

单片机蜂鸣器源代码+仿真

(^o^)-☆°.\*.°: 确实可以响，但是有两行看不懂，然后我挪到我的仿真...

嵌入式知识笔记

lswfcsdn: 开发板打印Rebooting in 5 seconds会一直重启啊 博主没出现重启的情况...

您愿意向朋友推荐“博客详情页”吗？



强烈不推荐



不推荐



一般般



推荐



强烈推荐

### 最新文章

用go语言写一个超级账本的智能合约（示例）

区块链超级账本hyperledger-fabric环境搭建，超详细

农业工程与信息技术介绍

2023年 4篇

2022年 64篇

2021年 92篇

2020年 37篇

2019年 97篇