

PHP汉字（UTF-8）转unicode码

作者：有故事的人 来源：东坡网

本文原地址：<http://www.dp1037.com/dpinfo-7-32-0.html>

东坡网，为帝国cms加油

最近要写一个与汉语字典有关的程序，发现许多生僻字，在页面不能显示，只能以unicode码的10进制实体来显示。于是，php将汉字转unicode码就成了迫切需求。用搜索找到好多个函数，可惜在试用后都放弃了，这些函数只支持普通汉字的转码，对于生僻字不支持。只好自己写一个，测试成功了。这个函数是根据unicode的二进制规则来写的。

```
function dp_HzToUnicode($hz){  
    $r=array();  
    $hex=str_replace("%","",urlencode($hz));  
    $r['hex']=$hex;  
    $bin2=base_convert($hex,16,2);  
    $dec=base_convert($hex,16,10);  
    if($dec<127){  
        $r['dec']=$dec;  
    }  
}
```



```
$r['unidoce']=$r['dec'];  
$r['bin2']=$bin2;  
return $r;  
}  
$bit=strlen($bin2)/8;  
$br=str_split($bin2,8);  
$true_bin2='';  
for($i=0;$i<$bit;$i++){  
    if($i==0){  
        $true_bin2.=substr($br[$i],$bit+1);  
    }else{  
        $true_bin2.=substr($br[$i],2);  
    }  
}  
$r['bin2']=$true_bin2;  
$unidoce=base_convert($true_bin2,2,16);  
$r['unidoce']=$unidoce;  
$html=base_convert($true_bin2,2,10);
```

```
$r['dec']=$html;  
return $r;  
}
```

更多 建站技术文档 请访问 <http://www.dp1037.com/dpclass-7-0/>

文章生成doc功能，由[东坡网](http://dp1037.com)开发