

# Unicode 字符编码表|汉字Unicode编码的区间为:

## 0x4E00→0x9FA5(转)

十进制		十六进制		字符数	编码分类 (中文)	编码分类 (英文)
起始	终止	起始	终止	(个)		
0	127	0000	007F	128	C0控制符及基本拉丁文	C0 Control and Basic Latin
128	255	0080	00FF	128	C1控制符及拉丁文补充-1	C1 Control and Latin 1 Supplement
256	383	0100	017F	128	拉丁文扩展-A	Latin Extended-A
384	591	0180	024F	208	拉丁文扩展-B	Latin Extended-B
592	687	0250	02AF	96	国际音标扩展	IPA Extensions
688	767	02B0	02FF	80	空白修饰字母	Spacing Modifiers
768	879	0300	036F	112	结合用读音符号	Combining Diacritics Marks
880	1023	0370	03FF	144	希腊文及科普特文	Greek and Coptic
1024	1279	0400	04FF	256	西里尔字母	Cyrillic
1280	1327	0500	052F	48	西里尔字母补充	Cyrillic Supplement
1328	1423	0530	058F	96	亚美尼亚语	Armenian
1424	1535	0590	05FF	112	希伯来文	Hebrew
1536	1791	0600	06FF	256	阿拉伯文	Arabic
1792	1871	0700	074F	80	叙利亚文	Syriac
1872	1919	0750	077F	48	阿拉伯文补充	Arabic Supplement
1920	1983	0780	07BF	64	马尔代夫语	Thaana
1984	2047	07C0	07FF	64	西非书面语言	N'Ko
2048	2143	0800	085F	96	阿维斯塔语及巴列维语	Avestan and Pahlavi
2144	2175	0860	087F	32	Mandaic	Mandaic
2176	2223	0880	08AF	48	撒马利亚语	Samaritan
2304	2431	0900	097F	128	天城文书	Devanagari
2432	2559	0980	09FF	128	孟加拉语	Bengali
2560	2687	0A00	0A7F	128	锡克教文	Gurmukhi
2688	2815	0A80	0AFF	128	古吉拉特文	Gujarati
2816	2943	0B00	0B7F	128	奥里亚文	Oriya

2944	3071	0B80	0BFF	128	泰米尔文	Tamil
3072	3199	0C00	0C7F	128	泰卢固文	Telugu
3200	3327	0C80	0CFF	128	卡纳达文	Kannada
3328	3455	0D00	0D7F	128	德拉维族语	Malayalam
3456	3583	0D80	0DFF	128	僧伽罗语	Sinhala
3584	3711	0E00	0E7F	128	泰文	Thai
3712	3839	0E80	0EFF	128	老挝文	Lao
3840	4095	0F00	0FFF	256	藏文	Tibetan
4096	4255	1000	109F	160	缅甸语	Myanmar
4256	4351	10A0	10FF	96	格鲁吉亚语	Georgian
4352	4607	1100	11FF	256	朝鲜文	Hangul Jamo
4608	4991	1200	137F	384	埃塞俄比亚语	Ethiopic
4992	5023	1380	139F	32	埃塞俄比亚语补充	Ethiopic Supplement
5024	5119	13A0	13FF	96	切罗基语	Cherokee
5120	5759	1400	167F	640	统一加拿大土著语音节	Unified Canadian Aboriginal S
5760	5791	1680	169F	32	欧甘字母	Ogham
5792	5887	16A0	16FF	96	如尼文	Runic
5888	5919	1700	171F	32	塔加拉语	Tagalog
5920	5951	1720	173F	32	Hanunóo	Hanunóo
5952	5983	1740	175F	32	Buhid	Buhid
5984	6015	1760	177F	32	Tagbanwa	Tagbanwa
6016	6143	1780	17FF	128	高棉语	Khmer
6144	6319	1800	18AF	176	蒙古文	Mongolian
6320	6399	18B0	18FF	80	Cham	Cham
6400	6479	1900	194F	80	Limbu	Limbu
6480	6527	1950	197F	48	德宏泰语	Tai Le
6528	6623	1980	19DF	96	新傣仂语	New Tai Lue
6624	6655	19E0	19FF	32	高棉语记号	Kmer Symbols
6656	6687	1A00	1A1F	32	Buginese	Buginese
6688	6751	1A20	1A5F	64	Batak	Batak
6784	6895	1A80	1AEF	112	Lanna	Lanna
6912	7039	1B00	1B7F	128	巴厘语	Balinese
7040	7088	1B80	1BB0	49	巽他语	Sundanese

7104	7167	1BC0	1BFF	64	Pahawh Hmong	Pahawh Hmong
7168	7247	1C00	1C4F	80	雷布查语	Lepcha
7248	7295	1C50	1C7F	48	Oi Chiki	Oi Chiki
7296	7391	1C80	1CDF	96	曼尼普尔语	Meithei/Manipuri
7424	7551	1D00	1D7F	128	语音学扩展	Phonetic Extensions
7552	7615	1D80	1DBF	64	语音学扩展补充	Phonetic Extensions Supplement
7616	7679	1DC0	1DFF	64	结合用读音符号补充	Combining Diacritics Marks Supplement
7680	7935	1E00	1EFF	256	拉丁文扩充附加	Latin Extended Additional
7936	8191	1F00	1FFF	256	希腊语扩充	Greek Extended
8192	8303	2000	206F	112	常用标点	General Punctuation
8304	8351	2070	209F	48	上标及下标	Superscripts and Subscripts
8352	8399	20A0	20CF	48	货币符号	Currency Symbols
8400	8447	20D0	20FF	48	组合用记号	Combining Diacritics Marks for Symbols
8448	8527	2100	214F	80	字母式符号	Letterlike Symbols
8528	8591	2150	218F	64	数字形式	Number Form
8592	8703	2190	21FF	112	箭头	Arrows
8704	8959	2200	22FF	256	数学运算符	Mathematical Operator
8960	9215	2300	23FF	256	杂项工业符号	Miscellaneous Technical
9216	9279	2400	243F	64	控制图片	Control Pictures
9280	9311	2440	245F	32	光学识别符	Optical Character Recognition
9312	9471	2460	24FF	160	封闭式字母数字	Enclosed Alphanumerics
9472	9599	2500	257F	128	制表符	Box Drawing
9600	9631	2580	259F	32	方块元素	Block Element
9632	9727	25A0	25FF	96	几何图形	Geometric Shapes
9728	9983	2600	26FF	256	杂项符号	Miscellaneous Symbols
9984	10175	2700	27BF	192	印刷符号	Dingbats
10176	10223	27C0	27EF	48	杂项数学符号-A	Miscellaneous Mathematical Symbols-A
10224	10239	27F0	27FF	16	追加箭头-A	Supplemental Arrows-A
10240	10495	2800	28FF	256	盲文点字模型	Braille Patterns
10496	10623	2900	297F	128	追加箭头-B	Supplemental Arrows-B
10624	10751	2980	29FF	128	杂项数学符号-B	Miscellaneous Mathematical Symbols-B
10752	11007	2A00	2AFF	256	追加数学运算符	Supplemental Mathematical Operators
11008	11263	2B00	2BFF	256	杂项符号和箭头	Miscellaneous Symbols and Arrows

11264	11359	2C00	2C5F	96	格拉哥里字母	Glagolitic
11360	11391	2C60	2C7F	32	拉丁文扩展-C	Latin Extended-C
11392	11519	2C80	2CFF	128	古埃及语	Coptic
11520	11567	2D00	2D2F	48	格鲁吉亚语补充	Georgian Supplement
11568	11647	2D30	2D7F	80	提非纳文	Tifinagh
11648	11743	2D80	2DDF	96	埃塞俄比亚语扩展	Ethiopic Extended
11776	11903	2E00	2E7F	128	追加标点	Supplemental Punctuation
11904	12031	2E80	2EFF	128	CJK 部首补充	CJK Radicals Supplement
12032	12255	2F00	2FDF	224	康熙字典部首	Kangxi Radicals
12272	12287	2FF0	2FFF	16	表意文字描述符	Ideographic Description Characters
12288	12351	3000	303F	64	CJK 符号和标点	CJK Symbols and Punctuation
12352	12447	3040	309F	96	日文平假名	Hiragana
12448	12543	30A0	30FF	96	日文片假名	Katakana
12544	12591	3100	312F	48	注音字母	Bopomofo
12592	12687	3130	318F	96	朝鲜文兼容字母	Hangul Compatibility Jamo
12688	12703	3190	319F	16	象形字注释标志	Kanbun
12704	12735	31A0	31BF	32	注音字母扩展	Bopomofo Extended
12736	12783	31C0	31EF	48	CJK 笔画	CJK Strokes
12784	12799	31F0	31FF	16	日文片假名语音扩展	Katakana Phonetic Extensions
12800	13055	3200	32FF	256	封闭式 CJK 文字和月份	Enclosed CJK Letters and Months
13056	13311	3300	33FF	256	CJK 兼容	CJK Compatibility
13312	19903	3400	4DBF	6592	CJK 统一表意符号扩展 A	CJK Unified Ideographs Extension A
19904	19967	4DC0	4DFF	64	易经六十四卦符号	Yijing Hexagrams Symbols
19968	40895	4E00	9FBF	20928	CJK 统一表意符号	CJK Unified Ideographs
40960	42127	A000	A48F	1168	彝文音节	Yi Syllables
42128	42191	A490	A4CF	64	彝文字根	Yi Radicals
42240	42527	A500	A61F	288	Vai	Vai
42592	42751	A660	A6FF	160	统一加拿大土著语音节补充	Unified Canadian Aboriginal Syllables Supplement
42752	42783	A700	A71F	32	声调修饰字母	Modifier Tone Letters
42784	43007	A720	A7FF	224	拉丁文扩展-D	Latin Extended-D
43008	43055	A800	A82F	48	Syloiti Nagri	Syloiti Nagri

43072	43135	A840	A87F	64	八思巴字	Phags-pa
43136	43231	A880	A8DF	96	Saurashtra	Saurashtra
43264	43391	A900	A97F	128	爪哇语	Javanese
43392	43487	A980	A9DF	96	Chakma	Chakma
43520	43583	AA00	AA3F	64	Varang Kshiti	Varang Kshiti
43584	43631	AA40	AA6F	48	Sorang Sompeng	Sorang Sompeng
43648	43743	AA80	AADF	96	Newari	Newari
43776	43871	AB00	AB5F	96	越南傣语	Vi?t Thái
43904	43936	AB80	ABA0	33	Kayah Li	Kayah Li
44032	55215	AC00	D7AF	11184	朝鲜文音节	Hangul Syllables
55296	56319	D800	DBFF	1024	High-half zone of UTF-16	High-half zone of UTF-16
56320	57343	DC00	DFFF	1024	Low-half zone of UTF-16	Low-half zone of UTF-16
57344	63743	E000	F8FF	6400	自行使用区域	Private Use Zone
63744	64255	F900	FAFF	512	CJK 兼容象形文字	CJK Compatibility Ideographs
64256	64335	FB00	FB4F	80	字母表达形式	Alphabetic Presentation Form
64336	65023	FB50	FDFE	688	阿拉伯表达形式A	Arabic Presentation Form-A
65024	65039	FE00	FE0F	16	变量选择符	Variation Selector
65040	65055	FE10	FE1F	16	竖排形式	Vertical Forms
65056	65071	FE20	FE2F	16	组合用半符号	Combining Half Marks
65072	65103	FE30	FE4F	32	CJK 兼容形式	CJK Compatibility Forms
65104	65135	FE50	FE6F	32	小型变体形式	Small Form Variants
65136	65279	FE70	FEFF	144	阿拉伯表达形式B	Arabic Presentation Form-B
65280	65519	FF00	FFEF	240	半型及全型形式	Halfwidth and Fullwidth Form
65520	65535	FFF0	FFFF	16	特殊	Specials

UTF-8有点类似于Haffman编码，它将Unicode编码为：  
0x00-0x7F的字符，用单个字节来表示；  
0x80-0x7FF的字符用两个字节表示；  
0x800-0xFFFF的字符用3字节表示；

汉字的unicode范围是： 0x4E00~0x9FA5

其实这个范围还包括了中，日，韩的字符。