

中华人民共和国国家标准

文字条目通用排序规则

GB/T 13418—92

Character and entry filing principles

本标准参照采用国际标准 ISO 7154—1983《书目编档规则》。

1 主题内容与适用范围

本标准规定了汉字字符和非汉字字符排序通用的基本规则。

本标准适用于各种名称(人名、团体名、国名、地名、题名、品名、物名、会议名等)、字词、目录、编号、代码、数字、年代等一切要求有序印刷、显示等的排序处理。不同业务领域,由于排序条目不同,排序项和排序段的设定也可以不一样,在使用这一规则时,可根据本标准和本行业的具体需要补充制订相应的排序细则。

2 引用标准

GB 1988 信息处理 信息交换用七位编码字符集

GB 2312 信息交换用汉字编码字符集 基本集

3 术语

3.1 字母顺序 alphabetical order

字母在字母表中的顺序。

3.2 排序条目 filing entry

最高一级排序单位,是一个有组织的需要排序的信息集合,如书目款目、字词及注释、人名录(包括姓名及其介绍信息)等。其中一定含有排序项。

3.3 排序项 filing area

第二级排序单位,直属于排序款目。它是排序款目中一个独立的参加排序的信息集合。如书目款目中的书名、作者、出版者、人名款目中的人名等。

3.4 排序段 filing section

第三级排序单位。它从属于排序项。如姓名中的姓和名就是姓名排序项的两个排序段。书名中的正题名、副题名及分册名也是书名的不同排序段。

3.5 排序字符 filing character

最低一级排序单位。它不能再细分。排序字符按类型划分,各个类型的排序字符均有特定的排序值。字母、数字、汉字等都是不同类型的排序字符。

3.6 排序单元 filing unit

在排序中定义的一个实体,为排序层次中的一部分。排序层次中不同级别的排序单元是由特定名称区分的。

3.7 非排序单元 non-filing unit

款目中的不参加排序的信息。

4 一般规则

4.1 排序信息的区分

作为一个排序款目中的每个独立的排序单元,都应该是可以辨认的。辨认的方法或是按排序单元在款目中的位置,或是按其确定的区分符号。

例:

目录组织规则:图书部分/北京图书馆编目部. ——北京 书目文献出版社 1984年5月

说明:这是简要描述一本图书的目录款目。如果仅按书名排序,则开头的书名就是第1个排序项。如果采取更多项目排序,除书名外,著作责任者、出版者、出版年也可做排序项,其间的区分,其书面或显示形式就是“/”,“.”,“——”,空格等。若是计算机处理,还要通过与这些区分符号相对应的字段、子字段标识符或这些数据存储的物理位置加以区分。

4.2 排序顺序

在一个排序款目中,各独立的排序单元是按一定的次序排列构成排序顺序,它不一定与款目中各数据元素的印刷或显示顺序一致。在各行业的排序细则中应指定排序顺序。

例:

4.1中示例的书目款目,如果按作者作为主标目加以排序,则作者就是第1排序项,第2排序项可选择书名,还可选择第3、4等排序项。

4.3 排序级别

完整的排序单位的数据结构规定自上至下的四层排序级别。

第一级 排序款目

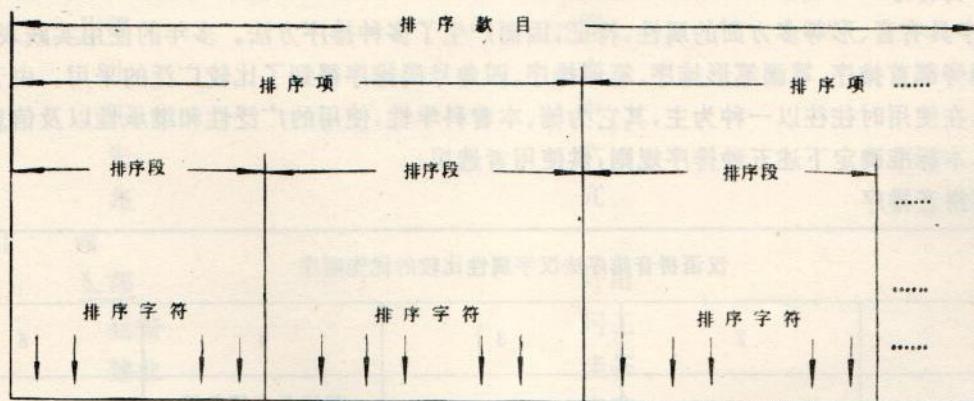
第二级 排序项

第三级 排序段

第四级 排序字符

排序级别可以少于四级。

具有四级层次排序顺序的构成



4.4 排序款目

4.4.1 排序款目的组成

根据本业务领域需要,确定排序款目的构成,设定排序项、排序段以及非排序项。

如文献目录款目,包括有:题名、责任者、出版发行项、载体形态、附注、标识号码、摘要、主题、分类等。其中题名、责任者、出版发行项、标识号码、主题、分类均可做排序项,其中一些排序项又包含排序段。

款目中的载体形态、附注、摘要是非排序项。

如机关团体通讯录款目,包括有:机关团体名称、通讯地址、邮政编码、电话号码等。其机关团体名称一定是排序项,通讯地址等是非排序项。

4.4.2 排序款目的排序

首先按排序款目的第一排序项比较。如果第一排序项相同时,则按第二排序项比较。如果第二排序项仍相同,则比较第三排序项,依此类推。

4.4.3 参见款目排序

在款目排序中,参见款目应与基本款目一起按相同的排序方法混合排序。在某一款目有参见条目时,也可将参见条目排在相关条目之后。

4.4.4 不同款目索引的排序

在某些业务领域,在基本款目排序之外,还需要抽取款目中某些项目编制索引,当超过一种以上索引时,则首先指定这些索引的先后顺序,然后实施索引本身排序。

4.5 排序项的排序

4.5.1 排序项的组成

排序项包含一个或一个以上的排序段。如姓名排序项,既包括姓,也包括名。如题名排序项,既包括正题名,也可能包括副题名、分册号、分册名。

4.5.2 排序项的排序

当对排序项进行排序处理时,要先指明参加排序的排序段的先后顺序,然后逐段实施排序。

如题名排序段的顺序指定为:正题名、分册号、分册名,排序时逐项比较。

中国人名排序,当姓在前,名在后,排序时可以以整个姓名为单位进行比较,也可严格依据姓与姓比较,名与名比较。

4.6 排序段

排序段已是排序的最基本的字符串,它由一个或多个字符组成。它可以包括有非排序的限定成分,但对限定成分必须做出明确标记。

对排序段实施排序处理,是对构成排序段的字符与字符的比较。

5 排序字符排序规则

5.1 汉字字符排序

由于汉字具有音、形等多方面的属性、特征,因而产生了多种排序方法。多年的使用实践表明,汉语拼音排序、偏旁部首排序、笔画笔形排序、笔形排序、四角号码排序得到了比较广泛的采用。由于这些方法各有所长,在使用时往往以一种为主,其它为辅。本着科学性,使用的广泛性和继承性以及信息处理技术的可能性,本标准确定下述五种排序规则,供使用者选用。

5.1.1 汉语拼音排序

汉语拼音排序法汉字属性比较的优先顺序

1	2	3	4	5
汉语拼音	音 调	总笔画数	起笔至末笔各笔 笔形序值	字符集编码

首先比较汉字的音,即按汉语拼音字母表顺序对汉字字符排列。如果拼音相同,比较音调,按阴平、阳平、上声、去声、轻声的次序对汉字字符排列。如果音和音调相同,比较汉字的总笔画数,从少到多对汉字字符排列。如果笔画数相同,比较汉字的起笔至末笔各笔笔形,依“横、竖、撇、点、折”顺序排列。若起笔至末笔各笔笔形仍相同,则按汉字在国家标准汉字编码字符集中的编码值从小到大排列。

例 1 字

ā	吖 阿 呵 啊 钢 脱
á	呵 啊 嘎
ă	呵 啊
à	呵 啊
ə	阿 呵 啊

例 2 词

安身	背气
安神	背弃
安石榴	倍数
安适	辈数

5.1.2 偏旁部首排序

偏旁部首排序法汉字属性比较的优先顺序

1	2	3	4
部首序号	部首外汉字笔画数	部首外汉字起笔至末笔各笔笔形序值	字符集编码

首先按汉字所属的部首、偏旁入部，同部首、偏旁的字按部首外汉字笔画数从少到多排列。如果部首外汉字笔画数相同，按部首外汉字起笔至末笔各笔笔形的“横、竖、撇、点、折”顺序排列。如果上述各笔笔形仍然相同，则按汉字在国家标准汉字编码字符集中的编码值从小到大排列。

例 1 字

<u>匕部</u>	𠂇
<u>匕</u>	𠂇
<u>北</u>	<u>风部</u>
<u>旨</u>	凤
<u>鬯</u>	夙
<u>匙</u>	凰
<u>儿部</u>	<u>一部</u>
<u>儿</u>	亡
<u>兀</u>	卞
<u>先</u>	六
<u>兆</u>	亢

例 2 词

5.1.3 笔画笔形排序

笔画笔形排序法汉字属性比较的优先顺序

1 总笔画数	2 起笔至末笔各笔笔形序值	3 字符集编码
--------	---------------	---------

首先按汉字的笔画数从少到多排列。同笔画数的字按起笔至末笔各笔笔形“横、竖、撇、点、折”顺序排列。如果上述笔形仍相同，则按汉字在国家标准汉字编码字符集中的编码值从小到大排列。

例 1 字

一 夫	二 玉	三 未
二 十	三 未	四 击
三 厂	四 击	五 匡
四 丁		

例 2 词

丰收 寿命	王国 奉行	玉瓜 奉命
玉瓜 奉命	未节 奏章	未节 奏章
未节 奏章	匡正 耕耘	匡正 耕耘
匡正 耕耘	未耜	

5.1.4 笔形排序

笔画笔形排序法汉字属性比较的优先顺序

1 起笔至末笔各笔笔形序值	2 字符集编码
---------------	---------

首先按汉字的起笔至末笔各笔笔形“横、竖、撇、点、折”次序对汉字排列。如果上述笔形相同，按汉字在国家标准汉字编码字符集中的编码值从小到大排列。

例 1 字

一 耗	二 粑	三 寿
二 粑	三 寿	四 烹
三 烹	丰 煮	五 奉
丰 煮	王 奉	六 禾
王 奉	末 禾	七 曷
末 禾	耕 曷	八 曷
耕 曷	杪	

例 2 词

芋头 水源	艾子 朽木	芍药 朴实
艾子 朽木	芍药 朴实	节支 机械
芍药 朴实	节支 机械	

术语
札结

权力

5.1.5 四角号码排序

四角号码排序法汉字属性比较的优先顺序

1 四角号码	2 横笔笔画数	3 总笔画数	4 起笔至末笔各笔笔形序值	5 字符集编码
-----------	------------	-----------	------------------	------------

首先比较汉字的四角号码,从小到大排列。同号码的字需区分字中“横”笔的多少,从少到多排列。“横”笔数相同要比较整字的总笔画数,从少到多排列。总笔画数相同,比较起笔至末笔各笔笔形,按“横、竖、撇、点、折”次序对汉字排列。如上述笔形仍相同,则按汉字在国家标准汉字编码字符集中的编码值从小到大排列。

例 1 字

玄 (00732)	亥 (00802)
衣 (00732)	奕 (00804)
充 (00732)	禦 (00901)
峦 (00772)	柰 (00904)
六 (00800)	亲 (00904)

例 2 词

市场	夜以继日
帝国	库存
离散数学	店员
亦步亦趋	唐三彩
庶民	

5.1.6 序数词排序

当一、二、三、……,壹、贰、叁、……这些汉字作为序数词使用,并要求按照其数值顺序排列时,则应按这些汉字本身所表示的数值从小到大加以排列。否则,不按所表示的数值排序。

例 1

北京一中
北京二中
北京三中
⋮
⋮

例 2

第一届环境科学国际研讨会
第二届环境科学国际研讨会
第三届环境科学国际研讨会
⋮
⋮

5.2 非汉字符串排序

5.2.1 数字

按数字所表示的数值从小到大排列。

例

8.15, 20, 96, 545, 620,

5.2.2 年代

按年代先后顺序排序。如公元前(B.C.)年代,应从其大数值到小数值排列;公元后(A.D.)年代应从小数值到大数值排列。

5.2.3 号码

5.2.3.1 数字型

按数值从小到大排列。

例 101, 102, 103,

对于从左到右按每位数字定义的非固定长号码,要从左到右按位比较其数值,从小到大排列。

例

426

43

432.02

44

441.289

5.2.3.2 字母型

按字母表中字母顺序排列。

例 1

BA

BB

BC

BD

BE

例 2

Aa

Ab

Ac

Ad

Ae

5.2.3.3 数字、字母混合型字符串

根据 GB 1988 字母表顺序逐位比较,排列。

例 1

AA01

AA02

AB01

AB02

⋮

⋮

AZ01

AZ02

⋮

⋮

例 2

1A

2A

2B

3B1

4B

5D

5D6

⋮

⋮

⋮

⋮

对于非十进制(如采用 16 进制)编码,则按其编码所表示的数值从小到大排列。

例

D4A8

D4A9

D4AA

D4AB

D4AC

D4AD

Colon

Colonel

5.2.5.2 日文假名

平假名(包括浊音、半浊音及拗音、促音等小写假名)83个,片假名(包括浊音、半浊音及拗音、促音等小写假名)86个,按照国家标准GB 2312所规定的次序排列。

5.2.5.3 希腊文字母

按大写 ΑΒΓ……Ω、小写 αβγ……ω 的顺序排列。

5.2.5.4 基里尔语系字母

按大写 АБВГ……Я、小写 абвг……я 的顺序排列。

例 1

Камов врачебной науки

Каннада-русский словарь

Кант

Кант и проблема знания

Кантабиле и вальс

Кануны

5.2.5.5 汉语拼音符号

汉语拼音符号虽属汉语语系,但仍按大写 ABCD……Z 和小写 abcd……z 的顺序排列。其中加声调的符号按符号本身的字顺排列。声调次序体现在字的拼音符号串的排序中。

例

chēn

chēng

chéng

chěng

chèng

chī

chí

chǐ

chì

chōng

:

:

对于完全以汉语拼音字符书写的款目,以词为单位(单音节词或多音节词),按拉丁字母表的顺序,逐位字符加以比较,决定款目顺序。这种排序并不考虑每个词所表示的汉字的其他属性。

例:

chai (柴)

chaicao (柴草)

chaiyou (柴油)

chaiyouji (柴油机)

chanpin (产品)

chanye (产业)

5.2.6 单词

在众多情况下,非汉字的文字信息的排序都是以单词为单位(即字符串形式)加以比较而进行的。在

单词范围内,通过逐位比较字母次序,决定单词的次序。而单词又是以空格为区分逐词进行比较。

5.2.6.1 缩写、简称与全称

无论缩写或全称,均以字面形式加以比较排列。缩写或简称中的圆点“·”,不列为排序因素,全部缩写字母作为一个字符串处理。

例 1

Iceland

I. E. E

IEEE

International

例 2

АО АН СССР

ДВЖД

ДВО АН СССР

Дружба народов

У политической карты мира

5.2.6.2 姓和名

书面形式应是姓在前,名在后。姓与姓的字顺相比,名与名的字顺相比。

例:

Heine, Heinrich.

Heine, William C.

带有前缀的姓,前缀与后面的姓要合成一个字符串进行比较、排列。

例:

D'Arcy, Albert

Defoe, James

De Francis, J.

Defrancis, John F.

De Garmo, Charles.

:

:

Ocantus

O'Casey, Sean

Vancouver

Van Dyke, Ernesh

Zur Mullen, Friedrich

Zu Tavern, B.

尊称 Mr. , Mrs. , Dr. , Sir 等不参加排序。

5.2.6.3 冠词、介词

原则上作为非排序单元处理。

5.2.6.4 序词、数词

按序词和数词的字面表达形式排序。

5.2.6.5 复合词

无论中间有否短横“-”,均按一个词排序。

例如:

South-east 按 Southeast 排

North-east 按 Northeast 排

Hand-book 按 Handbook 排

5.2.6.6 个别字母

凡是信息处理设备已收入这类字母图形及编码的,一律按其编码值排序。

对于信息处理设备未收入这类字母图形及编码的,可按传统办法处理。

例如:

ä,ć,ā,ë,ø,ü 等按 a,c,a,e,o,u 排。

ç,č,ł,ń,š,ł,按 c,c,i,n,s,t 排。

德文的 ö,ä,ü,有两种排法:

按 o,a,u 排。

按发音 ou,ae,ue 排。

5.2.7 标点及其他符号

一般情况下,标点符号属于非排序单元。

5.3 汉字字符与非汉字字符混合出现时排序

为了便于信息处理,依据国家标准 GB 2312中字符的排列顺序,确定以下排序的先后次序:

空格—序号—阿拉伯数码—拉丁字母(大写、小写)—日文假名(平假名、片假名)—希腊字母—俄文字母—汉字

例1:

1985年年鉴

2000年

A300喷气飞机

COBOL 程序设计

PASCAL 语言

新しい日语

新青年

英语学习 ABC

英语学习辅导

例2:

100 Alphabets Publicitaires

100% American

\$ 100 bond news

100 chapel talks

Las 100 mas famosas novelas

100 ø on the dollar

The £ 100 wager

100 x Zeichnen und Malen

100 years an orphan

ΣΓΜΒΩΠΑ

РУССКО-КИТАЙСКИЙ СЛОВАРЬ

附录 A
汉字部首表
(参考件)

一画		三画		四画	
1	一 1	37	力 23	70	己 49
2	丨 2	38	厶 24	71	弓 50
3	丿 3	39	又 25	72	中 51
4	、 4	40	乚 26	73	少▲
5	乙 5	41	干 27	74	女 52
6	乚▲	42	工 28	75	飞 53
7	乚▲	43	土 29	76	飛▲
8	乚▲	44	士▲	77	小 54
9	乚▲	(+ 211)		78	少▲
二画		45	升 30	79	子 55
10	十 6	46	大 31	80	马 56
11	厂 7	47	尤 32	81	馬▲
12	厂▲	48	兀▲	(糸 214)	
13	厂 8	49	允▲	82	幺 57
14	卜 9	(扌 115)		83	𠂇 58
15	卜▲	50	寸 33	84	王 59
16	口 10	51	弋 34	85	玉▲
17	刀▲	52	口 35	86	无 60
18	八 11	53	口 36	87	无▲
19	ノ▲	54	巾 37	88	韦 61
20	人 12	55	山 38	89	韋▲
21	入▲	56	彳 39	(攴 176)	
22	亻▲	57	乡 40	90	木 62
23	匚 13	(犭 93)		91	支 63
24	匕 14	58	夕 41	92	犬 64
25	儿 15	59	乚 42	93	犭▲
26	几 16	(匚 265)		94	刂 65
27	几▲	60	彳 43	95	步▲
28	士 17	61	爿▲	96	车 66
29	冂 18	62	广 44	97	車▲
	(宀 240)	(宀 140)		98	戈 67
30	乚 19	63	门 45	99	比 68
31	匚 20	64	門▲	100	牙 69
32	匚 21	(宀 145)		101	瓦 70
33	巳▲	65	宀 46	102	止 71
	(阝(左)252)	(宀 232)		103	支 72
	(阝(右)229)	66	丂 47	104	乚▲
34	刀 22	67	曰▲	105	日 73
35	乚▲	68	匚▲	106	曰▲
36	匚▲	69	尸 48	107	曰▲

146 氵▲	181 而 125	224 阝▲	261 骨 182
(水 207)	182 頁 126	(見 109)	(韋 89)
五画	183 頁▲	(見 111)	262 香 183
(玉 85)	184 至 127	225 里 156	263 鬼 184
147 示 98	185 丂 128	226 足 157	264 食 185
148 丷▲	186 虎▲	227 𠂇▲	265 𠂇▲
149 甘 99	187 虫 129	228 邑 158	(風 131)
150 石 100	188 肉 130	229 阝(右)▲	266 音 186
151 龙 101	189 缶 131	230 身 159	266 首 187
152 龍▲	190 舌 132	231 趾 160	(飛 76)
(夕 95)	191 竹 133	232 辵▲	十画
153 业 102	192 𠂇▲	233 采 161	268 禄 188
154 目 103	193 白 134	234 谷 162	269 彩 189
155 田 104	194 自 135	235 犭 163	(馬 81)
156 四 105	195 血 136	236 龜 164	270 門 190
157 皿 106	196 舟 137	237 龜▲	271 高 191
(钅 254)	197 色 138	238 角 165	十一画
158 生 107	198 齐 139	239 言 166	272 黃 192
159 矢 108	199 齐▲	240 讠▲	(麥 216)
160 禾 109	200 衣 140	241 辛 167	(爾 224)
161 白 110	201 丷▲	八画	(烏 164)
162 瓜 111	202 羊 141	242 青 168	(魚 256)
163 鸟 112	203 羊▲	(長 119)	273 麻 193
164 烏▲	204 羊▲	243 草 169	274 鹿 194
165 扌 113	205 米 142	244 雨 170	十二画
166 立 114	206 車 143	245 非 171	275 鼎 195
167 穴 115	207 車▲	246 齒 172	276 黑 196
(圭 208)	208 圭▲	247 齒▲	277 委 197
168 宀 116	209 艮 144	(虎 186)	十三画
169 正▲	210 卯 145	248 龜 173	278 鼓 198
(辵 201)	211 丂▲	249 眚▲	(鼴 249)
170 皮 117	212 羽 146	250 佳 174	279 鼠 199
171 犭 118	213 糸 147	251 阜 175	十四画
(水 146)	214 纟▲	252 阝(左)▲	280 鼻 200
172 矛 119	七画	(門 64)	(齊 199)
(母 143)	215 麦 148	253 金 176	十五画
六画	216 麥▲	254 钅▲	(齒 247)
173 未 120	(長 120)	255 魚 177	十六画
174 耳 121	217 走 149	256 魚▲	(龍 152)
175 老 122	218 赤 150	257 隶 178	十七画
176 丂▲	(車 97)	九画	(龜 237)
177 臣 123	219 豆 151	258 革 179	281 爪 201
178 西 124	220 西 152	(頁 183)	(龜 237)
179 西▲	221 辰 153	259 面 180	
180 西▲	222 犀 154	260 齉 181	
	223 齒 155		

注：① 每个部首左边数字为281个部首的序号。

每个部首右边数字为201个部首的序号。

② 有的部首右侧注有▲，表示该部首是辅形部首。

③ 括号()内部首，是便于查找的按笔画排列的参见部首。

④ 该部首是由中国文字改革委员会于80年代初期组织三十多个单位制订的。现已被《汉语大辞典》、《汉语大字典》、《国家经济信息系统规范与标准化设计》、《汉字属性字典》等采用。

附加说明：本标准由全国文献工作标准化技术委员会第四分会提出并负责起草。本标准主要起草人朱岩、朱南。

本标准由全国文献工作标准化技术委员会第四分会提出并负责起草。本标准主要起草人朱岩、朱南。

2 引用标准

下列标准所包含的条文，通过在本标准中引用而构成本标准的条文。本标准的编辑人员在本标准的编辑过程中未将有关条款摘录或转述，但所有这些条款仍然是本标准不可分割的一部分。本标准与GB2312-80相比，除增加“次常用级”外，其余部分均无变化。

3 词谱集结构

本标准共收入词谱4449条，按使用频率分三类：第一类（最常用级）1733条；第二类（常用级）1730条；第三类（次常用级）13256条。

本标准收入的词谱以二字、三字、四字词语为主，所收词谱的最大长度为12个汉字。

4 收词原则

本标准根据GB2312，并参照《汉语拼音正词法基本规则》和《信息处理用现代汉语分词规范》，收集书面语言中使用频率高以及日常通用的词汇和结合紧密、使用稳定的短语（参见附录A）。

5 词谱集

本标准的词谱按第一类（最常用级）、第二类（常用级）、第三类（次常用级）的顺序排列，每类又按二字、三字、四字和四字以上词谱表并按GB2312字符的顺序排列。

5.1 一、二字词谱集（第一类（最常用级））

5.1.1 第一类二字词谱表

爱戴	爱房	爱好	爱护	爱民	爱情	安定	安插	安顿	安培	安全	安危
安慰	安稳	安心	安详	安放	安顿	按时	按部	按行	按图	按语	惯性
关门	八十	八一	八月	宠爱	宠爱	推崇	推崇	推崇	推崇	百货	百万
百姓	摆布	摆设	摆布	摆设	摆布	摆手	摆手	摆手	摆手	搬运	搬运
朋友	颁奖	版本	遍地	半斤	半月						
办法	办厂	办法	办公	办好	办通	办通	办通	办通	办通	办法	办法