

## 第四章 五笔字型编码法

对于从未学过五笔字型的人士，我们建议最好“一步到位”直接学习新世纪版五笔字型，因为新世纪版（第三代）是五笔字型编码技术的“终极版”，其中的字根和字根键位以后不再改动了。

对于原来已学会 86 版、98 版的人士，只要仔细看一看第四章第六节（前页）的新世纪版与 86 版字根变动表，也不难自然而然地过渡到新世纪版。原因是，新世纪版完全符合规范、更加好记好用，三个版本的字、词“取码法”及识别码，也都是相同的。

一些人士认为五笔字型“难学”，这是缺乏分析的。

**学习起来并不难。**只要按规律记住 25 个字母键上的 125 种字根，即**字根键位图**（可以背一背助记歌帮助记忆），再按照**编码流程图**所示的规则拆字和取码，就能很快学会。

其实，用五笔字型打字，就等于在键盘上用“字根”写字，只不过是按一下键，就等于“写了一个字根”罢了。只要会写字，一两天学会编码并不难，真正难的是打字速度！因为多数中国人没有从小练好英文打字指法，十个指头不听话，年龄越大，越不听话！怎能打字？但只要方法得当又常加练习，每分钟输入 50~80 个字是不难的。

**学一阵子，用一辈子！**

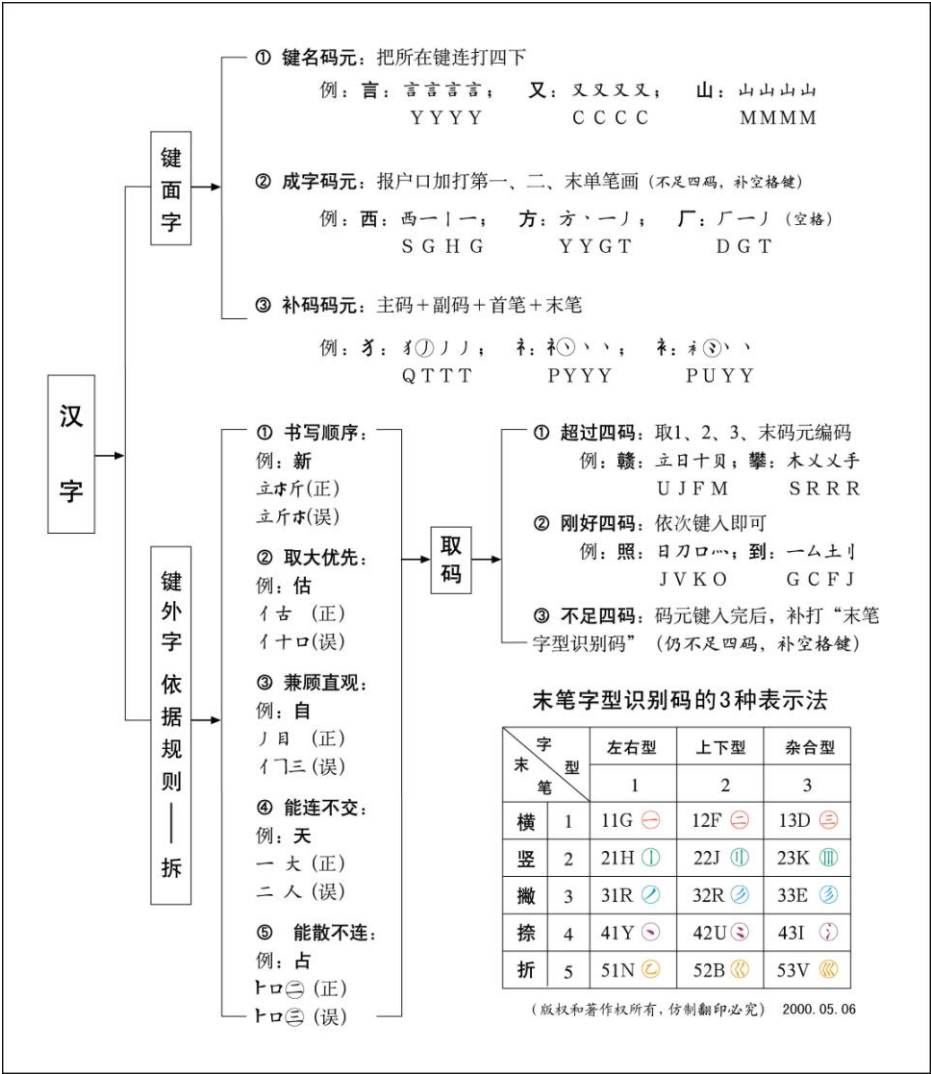
上一章掌握了字根所在的区和键位，下面学习拆字和取码法。



第一节 编码流程图

五笔字型将成千上万个汉字首先分成两大类：键面上有的“键面字”和键面上没有的“键外字”。两类汉字的取码法按以下流程图分别取码。

五笔字型编码流程图



## 第二节 “键面字”的编码输入法

五笔字型字根键盘上，本身是汉字的字根，叫“键面字”。“键面字”分为三类，其输入法分别是：

### 1、键名字的输入

每一个键位上，最左上角的那个黑体字的码元，叫键名字，它是“一键之名”。以 G 键为例，其左上角的字根“王”便是“键名字”或“键名”。



键名字“王”；



键名字“言”。

“键名字”的输入法是：把所在的键连打 4 下。例如：

1 区 1 位键名：王 11 11 11 11 (GGGG)

3 区 2 位键名：白 32 32 32 32 (RRRR)

4 区 5 位键名：之 45 45 45 45 (PPPP)

5 区 3 位键名：女 53 53 53 53 (VVVV)

在五笔字型中，键名字共有如下 25 个：

1 区：王土大木工（对应键位：GFDSA）

2 区：目日口田山（对应键位：HJKLM）

3 区：禾白月人金（对应键位：TREWQ）

4 区：言立水火之（对应键位：YUIOP）

5 区：己子女又乚（对应键位：NBVCX）

### 2、成字字根的输入






键面上除键名外，凡本身是汉字的码元，叫“成字字根”或“成字码元”。其输入法是：先打一下它所在的键（这一下俗称“报户口”），再打第一个、第二个，以及它的最末一个单笔画，最多 4 下，不足 4 下，补一个

空格键。例如：



成字根	报户口	第一笔	第二笔	末笔	输入法
五	五 (11 G)	一 (11 G)	丨 (21 H)	一 (11 G)	11 11 21 11 G G H G
古	古 (13 D)	一 (11 G)	丨 (21 H)	一 (11 G)	13 11 21 11 D G H G
车	车 (24 L)	一 (11 G)	乙 (51 N)	丨 (21 H)	24 11 51 21 L G N H
广	广 (44 O)	丶 (41 Y)	一 (11 G)	丿 (31 T)	44 41 11 31 O Y G T
寸	寸 (12 F)	一 (11 G)	丨 (21 H)	丶 (41 Y)	12 11 21 41 F G H Y
彳	彳 (43 I)	丶 (41 Y)	丶 (41 Y)	ㄟ (11 G)	43 41 41 11 I Y Y G
亻	亻 (34 W)	丿 (31 T)	丨 (21 H)	(空格)	34 31 21 W T H

当输入十、七、九、二、几、儿、乃、刀……这一类只有 2 个笔画的字根时，“报户口”之后就只能有两个笔画了，不足 4 个码，笔画打完，要再补打一下“空格键”表示结束。

3、补码码元的输入

在字根表中，用方圆框“框住”的、、、（只用于繁体字）共 4 个字根，是“补码码元”，它们作为字根参与编码时，像姓氏中的复姓诸葛、司马一样，要编 2 个码：“主码（即码元所在键位）+补码（规定取该码元最后的笔画结构）”。如下表所示：

补码码元	主码（第一码）	补码（第二码）
	 35. Q	 31. T
	 45. P	 41. Y
	 45. P	 42. U
	 32. W	 41. Y

提示：表中带圆圈的笔画、等，是“补码”的“笔画表示形式”，作为一个符号，用以提示编码。

这 4 个补码码元中的“𠂇、㇀、㇁”等三个字根，本身也是汉字，这三个汉字的编码规则是，要先“报户口”（主码+补码）（已占用 2 个码）、再打该字的第 1 笔和最后 1 笔，共取 4 码。即：

「报户口」

第1码 + 第2码 + 首笔画 + 末笔画

𠂇：𠂇 ① 丿 丿 （字根拆分）  
Q T T T （编码输入）

㇀：㇀ ⑤ 丶 丶 （字根拆分）  
P Y Y Y （编码输入）

㇁：㇁ ③ 丶 丶 （字根拆分）  
P U Y Y （编码输入）

不难发现，并不是键位上所有的字根都可以“打出来”！如“11”和补码码元“𠂇”等，那是因为这些字根本身并不是汉字。

**提示：**“补码码元”的设计，完全是为了保证键面字根的完整性和部件规范。其实，它们的编码效果和 86 版完全一样，只是换了个“说法”罢了！也就是说，对于𠂇 𠂇 ㇀，你就当作是 86 版的“丿”和“丶”仍然在 Q 和 P 键上，仍然按 86 版的𠂇=丿丿、㇀=丶丶、㇁=丶丶 编码输入，效果完全一样！

### 第三节 “键外字”的拆分法

本节是学习五笔字型的重点内容。

凡是“字根总表”上没有的汉字，都是“键外字”。它们都是由几个码元（字根）组合而成的，我们也把这类字称为“多元字”。

对于“多元字”，我们的办法，只有一个字——拆！

我们必须把由若干个字根构成的字，都“拆”成为一串字根表上列有的“字根”！怎么“拆”呢？必须按照一定的规则“拆”。科学、实用又没有“二义性”的“拆”法，共有以下 5 项规则（参见本章第一节“编码流程图”）：

#### 1、书写顺序

“合体字”“拆”成字根，大多数情况下，要按照正确的书写顺序进行。例如：

新：立 木 斤（顺序正确）

立 斤 木（顺序错误）

#### 2、取大优先

要“拆”出“笔画尽可能多”的“字根”。要以“再添一个笔画，便不能构成笔画更多的字根”为限度！例如：

失：乚 人（正确，再添一笔，乚就不是字根了）

㇏ 大（错误，再添一笔成为“乚”，比“㇏”要大）

估：亻 古（正确）

亻 十 口（错误，因为“口”可添到前面的“十”上，“凑”成更多笔画的字根“古”）

注：“取大优先”，俗称“尽量往前凑”。因为“向前凑”总是有限度的，要凑成字根表中笔画更多（更大）的一个！否则，便没有“章法”了，势必有许多种“拆法”甚至都要拆成单笔画！在汉字拆分中，这是一个最常用的、保证拆分唯一性的基本原则。

### 3、兼顾直观

在拆分汉字时，为了使拆出的字根特征明显易辨，有时就得暂时牺牲“书写顺序”和“取大优先”了，成为极为个别的例外情况。例如：

“国”字，按“书写顺序”，应拆成“冂王、一”。但这样拆法，不但有悖于“国”字的字源（“国”的外框“口”本来是城墙的意思），也不能使字根“口”直观易辨。于是，这种情况下，我们只好违背“书写顺序”，按“口王、”的顺序拆分了。

同样，不能按“取大优先”拆分的例子是：

自：丿 目 （正确，但并没有“取大优先”）

亻 冂 三 （虽按“取大优先”拆分了，但不直观，有悖于字源）

注：“自”是一个指事字，上面是一个手指，下面的“目”原本是鼻子的象形，意即用指头指着鼻子代表“自己”。

### 4、能连不交

请看以下取码实例：

天：一 大 （正确，“一”与“大”是相连接的关系，比较直观）

二 人 （错误，“二”与“人”交叉在一起了）

生：丿 圭 （正确，“丿”与“圭”是相连的，直观可取）

乚 土 （错误，“乚”与“土”是交叉的，此处不可取）

实际上，把两个字根当作相“连”，比当作相“交”更为“直观”，更能显现出字根的笔画结构特征，更易于辨认。“能连不交”，意思是既能按“连”又能按“交”的时候，要首先按“连”的关系对待，以此保证拆分的唯一性。

### 5、能散不连

有时候，一个汉字的几个字根，都不是单笔画，这些字根之间的关系，常常在“散”和“连”之间模棱两可。如：

占：卜 口，两个字根按“连接”处理，便是杂合型（3 型）字；如按

“离散”处理，便是上下型（2 型）字。

矢：𠂆 大，两个字根按“连接”处理，“矢”便是杂合型（3 型）字；两个字根如果按“离散”处理，“矢”便是上下型（2 型）字。

又如：羊：𠂆 𠂆（两个字根，可视为散，也可当作连）

午：𠂆 十（两个字根，可视为散，也可当作连）

当遇到少数这种既能当作“散”，又能当作“连”的情况时，为了避免拆分的二义性，我们规定：

只要不是单笔画（两个笔画或两个以上笔画），一律按“能散不连”判别；也就是在判定这类字的字型时，优先当作是“散”的关系。这一规定，在确定这类汉字的“识别码”时非常有用。

由此可见，以上两例中的“占”和“矢”，都被认为是“上下型”，即作为 2 型字（这样规定，仅仅是为取码的严密性设计，并不影响你的书写习惯和汉字结构！）。



## 第四节 四元字及多元字的取码法

### 一、“四元字”的编码规则

键外字中的“四元字”，是指刚好由四个字根构成的字。其拆分之后的取码方法是“依照书写顺序输入字根”。例如：

照：日 刀 口 灬  
22 53 23 44  
J V K O

荷：艹 亻 丁 口  
15 34 14 23  
A W S K

### 二、“多元字”的取码规则——“前三末一”（一二三末）

键外字中的“多元字”，是指由4个以上的字根构成的字。这种字，不管实际上能“拆”成几个字根，我们只需“按书写顺序，取拆分结果的第一二三及最末一个字根”便可，俗称“一二三末”，共输入四个码。例如：

憇：立 日 十 夕 工 贝 心  
42 22 12 51  
U J F N

耀：尢 一 儿 羽 亻 圭  
43 11 35 41  
I G Q Y

攀：木 乂 乂 木 大 手  
14 32 32 32  
S R R R

## 第五节 末笔字型识别码

汉字编码输入法的设计，要尽量减少重码，以提高输入的唯一性。但从以下两种情形我们看到，仅仅输入字根，很容易产生重码：

### 1、因构字的字根相同引起重码：

叭：口 八 (23 34 KW)      只：口 八 (23 34 KW)

### 2、因几个字根处在同一个键位，引起重码：



沐：氵 木 (43 14 IS)

汀：氵 丁 (43 14 IS)

洒：氵 西 (43 14 IS)

由以上两类例子可知，当遇到 2-3 个字根构成的汉字时，为了避免编码相同（重码），既有必要提取“字型信息”，又有必要从字根上“提取笔画特征信息”用于编码。复合这两种信息的一个附加码，就是“末笔字型识别码”简称“识别码”，“识别码”只追加在由 2-3 个字根构成的汉字编码中（见下节）。

“识别码”是由“末笔”代号加“字型”代号构成的一个“复合附加码”。1、2、3 型汉字的识别码共有 15 个（各有 3 种形式），其构成如下：

字 型 末 笔		左右型	上下型	杂合型
		1	2	3
横	1	11G ㄟ	12F ㄥ	13D ㄣ
竖	2	21H ㄴ	22J ㄴ	23K ㄴ
撇	3	31R ㄥ	32R ㄥ	33E ㄥ
捺	4	41Y ㄥ	42U ㄥ	43I ㄥ
折	5	51N ㄥ	52B ㄥ	53V ㄥ

例：

红：末笔 1，字型 1，  
识别码为 11（即 ㄟ）；

华：末笔 2，字型 2，  
识别码为 22（即 ㄴ）；

团：末笔 3，字型 3，  
识别码为 33（即 ㄣ）。

## 第六节 二元字及三元字的取码法

“键外字”中，只有 2 个字根的字，叫“二元字”；只有 3 个字根的字，叫“三元字”。输入时，键外字毫无例外地都要“拆”！

“二元字”或“三元字”的输入法是：

先“拆”成字根，输入字根后，再追加一个“末笔字型识别码”（简称“识别码”）。有了识别码可以大量减少重码。

“识别码”的简易直观表示法——用带圆圈的笔画表示“识别码”：

### 1、左右型（1 型）字的“识别码”

对于 1 型（左右型）字，字根输入之后，补打“1 个末笔画”，就等同于添加了“识别码”，用  $\ominus$  ① ② ③ ④ 表示之。例如：

江：氵 工  $\ominus$  （字根打完，补打 1 个末笔画“一”，相当于 11·G）

43 15 11

I A G

别：口 力 丨 ① （字根打完，补打 1 个末笔画“丨”，相当于 21·H）

23 33 22 21

K E J H

形：一 卅 彡 ② （字根打完，补打 1 个末笔画“丿”，相当于 31·T）

11 15 33 31

G A E T

### 2、上下型（2 型）字的“识别码”

对于 2 型（上下型）字，码元输入之后，补打“2 个末笔画”的字根，就等同于添加了“识别码”，用  $\ominus$  ⑪ ⑫ ⑬ ⑭ 表示之。例如：

字：宀 子  $\ominus$  （字根打完，补打 2 个末笔画“二”，相当于 12·F）

45 52 12

P B F

复：ㄣ 日 久 ㄣ (字根打完，补打 2 个末笔画“ㄣ”，相当于 42·U)

31 22 31 42

T J T U

花：艹 亻 匕 ㄣ (字根打完，补打 2 个末笔画“ㄣ”，相当于 52·B)

15 34 55 52

A W X B

### 3、杂合型（3 型）字的“识别码”

对于 3 型（杂合型）字，码元输完之后，补打“3 个末笔画”的字根，就等同于添加了“识别码”，用ㄣ ㄣ ㄣ ㄣ ㄣ 表示之。例如：

局：尸 丁 口 ㄣ (字根打完，补打 3 个末笔画“ㄣ”，相当于 13·D)

51 51 23 13

N N K D

朱：ㄣ 小 ㄣ (字根打完，补打 3 个点的末笔画如“ㄣ”，相当于 43·I)

32 43 43

R I I

远：二 儿 辶 ㄣ (字根打完，补打 3 个末笔画“ㄣ”，相当于 53·V)

12 35 45 53

F Q P V

注 1：凡是“包围型”的字，如全包围字“国、团”等，半包围字“这、庆”等，均以被包围的那个部分的“末笔”作为整个字的“末笔”来构成“识别码”，如“远”字，要以被包围的“儿”的末笔来构成“识别码”(53·V)。

注 2：识别码一共有 3 种表示法（请见 64 页下方的表），其编码的效果完全相同，都是同一个码！可以按照下面的方法打“识别码”，例如：

末笔横的 1 型字：打 11 键，就是 G 键，就是ㄣ（一个横）键！

末笔撇的 3 型字：打 33 键，就是 E 键，就是ㄣ（三个撇）键！

其所以如此，道理很简单：1 区 1 位（G）上有 1 个横ㄣ，3 区 3 位（E）上有 3 个撇ㄣ……

## 第七节 简码和容错码输入

### 一、简码输入

一些常用的字，除按它的“全码”可输入外，为减少打键次数，只输入其全码的最前边的 1 个、2 个或 3 个码，再加打空格键，也可以输入，这就是一、二、三级简码。简码可以提高输入效率。

“简码”正如人的名字，比如名字“赵丽梅”，就是“全码”，“二级简码”——丽梅，最亲近的人则叫“一级简码”——梅。

#### 1、一级简码（又称“高频字”）

将字母键按一下，再按一下空格键，即可打出 25 个最常用的汉字如下表：

区号 \ 位号	1	2	3	4	5
1	11G 一	12F 地	13D 在	14S 要	15A 工
2	21H 上	22J 是	23K 中	24L 国	25M 同
3	31T 和	32R 的	33E 有	34W 人	35Q 我
4	41Y 主	42U 产	43I 不	44O 为	45P 这
5	51N 民	52B 了	53V 发	54C 以	55X 经

#### 2、二级简码（只输入“全码”的前 2 个码）

化：亻 匕（34 55 WX）

李：木 子（14 52 SB）

理论上，二级简码共有  $25^2=625$  个，请见本书附录 1。

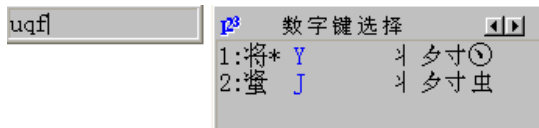
#### 3、三级简码（只输入“全码”的前 3 个码）：

想：木 目 心 (14 21 51, S H N)

巍：山 禾 女 (25 31 53, M T V)

三级简码的数量比较多,理论上 $25^3=15625$ 个,实际上有许多“空位”,这里不再列出。但输入软件对三级简码是作了标示的。在输入过程中会发现,“提示框”中有些字的后面带个“\*”号,就说明这样的字是简码字。

如:输入“将”,当输入前三个编码“uqf”后,我们看到“将”的后面有“\*”号,这就是告诉我们“将”字是有“三级简码”的,此时,不必打最后一个字根,只要按一下空格键,“将”字上屏。



注:上图中,“将”的后面还有一个字母“Y”,那是提示,还剩下一个编码“Y”没有输入,在有“\*”号的情况下,Y就不必输入,直接打空格键,三级简码字就可以上屏了。

有些字因为很常用,它可能有一级简码,也可能同时还有二级简码和三级简码。如“经”字,就有一、二、三级简码。

经: X (加空格)

经: X C (加空格)

经: X C G (加空格)



## 二、容错码

“容错码”的涵义是：“容易”编错，但“容许”按错码输入。例如：

面：丿 冂 丨 三 DMJD（正确，按笔顺取大优先）

丿 口 丨 二 DMJF（拆分容错）

左：ナ 工 ㇏ DAF（正确，视为上下型）

ナ 工 ㇏ DAD（字型容错）

如果按“容错码”输入，大一统软件会自动提示出正确的编码。例如对于“左”字输入“DAD”，在提示框中，会提示“左”的正确编码“DAF”，及正确的字根分解“ナ工㇏”（㇏是末笔为“一”的2型字的“识别码”）。



## 三、万能键 Z

输入汉字时，如果一时不知道某些字的编码，便可以用“万能键 Z”来代替“不知道的那个码”。Z 键的用途主要可分以下三种情况：

1. 当不知道某个字的拆分时，用 Z 代替不知道的字根，例：

键：𠂔 Z Z 又

2. 当不知道字根在哪个键位上时，用 Z 代替，例：

论：讠 人 Z Z

3. 当不知道字的“识别码”时，可用 Z 代替，例：

花：艹 匕 Z

万能键 Z 也叫**学习键**。一旦使用 Z 键，提示行中便会有比较多的字显示出来，其中会有你要的那个字，而且，字的后边还有正确码的提示。

当然，用 Z 键越多，提示行的字也就越多。如果你打了 4 个“Z”，那提示就把全部汉字都显示一遍——因为你没有告诉电脑，你要的那个字有什么特征。

## 第八节 词汇输入

在字母键上，打 4 个键，不用换档，既能打单字，又能打词汇，字、词之间没有界限，这是发明人 1983 年的一项重大创造——字词兼容。

在输入词汇时，不管多长的词汇，一律只打 4 下键，单字和词汇可以混合输入，字词之间不用任何换档或其它附加操作。词汇输入法为：

### 一、2 字词

取每个字“全码”的前两个码组成，共 4 码。例如：

经济：纟 ス 讠 文  
55 54 43 41 (XCIY)

建设：ㄣ 丰 讠 几  
53 11 41 34 (VGYW)

### 二、3 字词

前两个字，各取第一个码，最后一字取前两个码，共 4 码。例如：

电视机：日 礻 木 几  
22 45 14 25 (JPSW)

新加坡：立 力 土 尸  
42 33 12 21 (UEFH)

### 三、4 字以上的词

对于 4 个字或超过 4 个字的词，取第一、二、三及最后一个汉字的第一码，共 4 码。例如：

科学技术：禾 ㄣ 扌 木  
31 43 32 14 (TIRS)

常务委员会：ㄣ 久 禾 人  
43 31 31 34 (ITTW)

中华人民共和国：口 亻 人 口  
23 34 34 24 (KWWL)



第九节 王码符号双轨快捷输入法

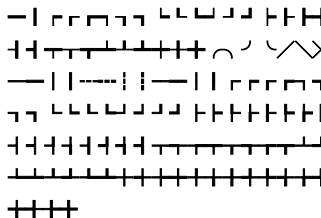
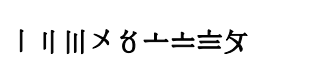


王码大一统软件 WBD2008 首创了“王码符号双轨快捷输入法”。

将常用符号分组，各起一个“汉字名称”，再用“五笔字型”或拼音（简拼）输入符号的“汉字名称”，便可快捷地输入 58 组计 900 个符号，成为大一统软件的一大亮点。符号的“汉字名称”及其输入法示例如下：

类别	序号	汉 字 名 称	86 版 编码	98 版 编码	新世纪 编码	拼音 输入	屏幕/提示行 显示内容
〔一〕 常 用 序 号	1	王码第一	GDTG	GDTG	GDTG	WMDY	1. (1) ① (→) i I 1
	2	王码第二	GDTF	GDTF	GDTF	WMDE	2. (2) ② (⇌) ii II 2
	3	王码第三	GDTD	GDTD	GDTD	WMDS	3. (3) ③ (⇨) iv III 3
	4	王码第四	GDTL	GDTL	GDTL	WMDS	4. (4) ④ (四) iv IV 4
	5	王码五六	GDGU	GDGU	GDGY	WMWL	5. (5) ⑤ (五) v V 5 6. (6) ⑥ (六) vi VI 6
	6	王码七八	GDAW	GDAW	GDAW	WMQB	7. (7) ⑦ (七) vii VII 7 8. (8) ⑧ (八) viii VIII 8
	7	王码九十	GDVF	GDVF	GDVF	WMJS	9. (9) ⑨ (九) ix IX 9 10. (10) ⑩ (十) x X 0
	8	十一十六	FGFU	FGFU	FGFY	SYSL	11. (11) XI 12. (12) XII 13. (13) 14. (14) 15. (15) 16. (16)
	9	十七二十	FAFF	FAFF	FAFF	SQES	17. (17) 18. (18) 19. (19) 20. (20)
	10	王码编号	GDXK	GDXK	GDXK	WMBH	N <sub>0</sub>
〔二〕	11	王码五星 STAR	GDGJ	GDGJ	GDGJ	WMWX	☆☆☆
	12	王码四方	GDLY	GDLY	GDLY	WMSF	◇◆□■
	13	王码三角	GDDQ	GDDQ	GDDQ	WMSJ	△▲▼▽▲▲▲▲
	14	王码圆圈	GDLL	GDLL	GDLL	WMYQ	○●◎⊙⊕⊖
	15	王码箭头	GDTU	GDTU	GDTU	WMJT	→←↑↓↖↗↘↙

[illegible]

73

外 文 字 母	49	罗马小写	LCIP	LCIP	LCIP	LMXX	α β γ δ ε ζ η θ ι κ .....
	50	俄文大写	WYDP	WYDP	WYDP	EWDX	А Б В Г Д Е Ё Ж З И .....
	51	俄文小写	WYIP	WYIP	WYIP	EWXX	а б в г д е ё ж з и .....
	52	日文大写	JYDP	JYDP	JYDP	RWDX	ア ア イ イ ウ ウ エ エ オ オ .....
	53	日文小写	JYIP	JYIP	JYIP	RWXX	あ あ い い う う え え お お .....
	54	音调标号	UYSK	UYSK	UYSK	YDBH	āáǎǎēēēēīīīīōōōōūūūū úúúúûûûûûê&amrññhng
〔七〕 编 辑 符 号	55	表格符号	GSTK	GSTK	GSTK	BGFH	
	56	杭州数码	SYOD	SYOD	SYOD	HZSM	
	57	特殊符号	TGTK	CGTK	TGTK	TSFH	
	58	阶梯黑块	BSLF	BSLF	BSLF	JTHK	

(以上表格内容及符号的分类命名法, 版权所有, 未经许可不得抄用。)

