

五笔字型输入法教程

长安大学信息工程学院
屈立成

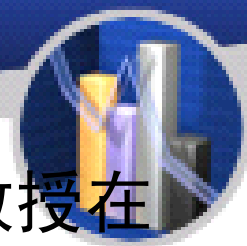


五笔字型教程



- 五笔字型输入法概述
- 五笔字型输入法的特点
- 汉字基本知识
- 五笔相关的名词解释
- 五笔字型的键盘设计
- 五笔字型字根分布的助记特征
- 五笔字型汉字拆分
- 五笔字型的汉字编码
- 五笔字型词组编码

五笔字型输入法概述



五笔字型是著名汉字信息处理专家王永民教授在五笔画基础上进一步完善的一种更效率的汉字输入方法，与其它音形类或纯音类输入法的一个不同点就是，它完全根据汉字的字形结构来进行编码，编码与一个汉字的读音没有任何关系。会五笔字型的操作员，即使碰到一个不会念的汉字，只要知道它怎样写，分成几部分，就可以将其输入计算机。

五笔字型”是一种高效率的汉字输入法，是只使用25个字母键，在键盘上汉字的笔画、字根为单位，向电脑输入汉字的方法。这一输入法，是在世界上占主导地位、应用最广的汉字键盘输入法。

五笔字型输入法的特点



1、世界上重码最少的汉字输入

纯形编码，编码的唯一性好。适合专职和非专职人员共同使用。按照形码设计三原理设计的字根键位分布，实现了同一键位上若干字根的相容性，合理地分配了编码空间，使重码减少到最低限度。平均每输入**10000**个汉字，仅有**1~2**个字需要人工挑选。

2、不受读音方言限制

中国方言多杂，南腔北调，加上一字多音，给拼音输入造成困难，五笔字型则不存在这些问题。

五笔字型输入法的特点



3、输入效率

“五笔字型”用双手10个手指打键。经过标准的指法训练，每分钟可向电脑中输入100个汉字。1998年“京城五笔字型大奖赛”，冠军的速度是在错1罚5的严厉评判规则下，每分钟输入生稿293个汉字。

4、字词兼容

用“五笔字型”既能输入单字，还能输入词汇。无论多么复杂的汉字最多只打4下键，无论多长的词汇也只打4下键。字与词之间，不要任何转换或附加操作，既符合汉字构词灵活、语句中字词很难切分的特点，又能大幅度地提高输入速度。

五笔字型输入法的特点



5、越打越顺

“五笔字型”依照“形码设计三原理”研究完成。
码元的组合符合“相容性”——使重码最少；
键位安排符合“规律性”——使字根易记易学；
指法设计的“协调性”——使得各个手指的击键负担趋于合理，打起来顺手，越打越顺手。

6、全球通用

“五笔字型”经过长久的大规模社会实践的检验，已成为在国内外占主导地位的汉字输入技术，具有很好的通用性。学会了“五笔字型”，到处都有现成的电脑可供使用；厂家的电脑类产品装入了“五笔字型”，全国乃至世界各地，都有成百上千的人不经训练便会操作。

汉字基本知识



1. 汉字的笔画

书写汉字时，一次写成的一个连续不断的线段成为笔画，从书写形态上来看汉字的笔画有：点、横、竖、撇、捺、挑（提）、钩、（左右）折共八种。

五笔字型中汉字的五种基本笔画是：

横（一）、竖（丨）、撇（丿）、捺（丶）、折（乙），

其中：

- （1）提笔“乚”实际是横“一”的变形。如：“现、场、特、扛”等，各字左部末笔的提 视为横。
- （2）点笔“丶”是捺“乚”的变形。如：“村、学、家、寸、冗”等，各字中的点 ，包括冂中的左点都为捺。
- （3）竖左钩“乚”是竖“丨”的变形。
- （4）一切带转折的笔画都归为折类。

汉字基本知识



2. 汉字的三个层次

笔画、部首、整字，是汉字结构的三个层次。

3. 汉字的书写顺序

先左后右，先上后下，先横后竖，先撇后捺，先内后外，先中间后两边，先进门后关门等。

4. 汉字的基本组成单位

汉字都是由**笔划**或**部首**组成的，一个汉字可以拆成一些最常用的基本单位，这个基本单位在五笔字型里就叫**字根**。

汉字基本知识



5. 汉字的基本结构

(1) 单体结构 由基本字根独立组成的汉字，

例如：目、一、手、山等。

(2) 左右结构 由左右两部分或左中右三部分构成，

例如：明、响、辩等。

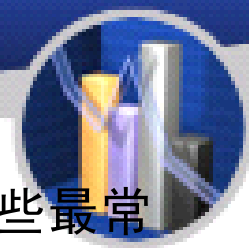
(3) 上下结构 由上下两部分或上中下三部分构成，

例如：昌、旦、莒、感等。

(4) 内外结构 由内外部分构成，

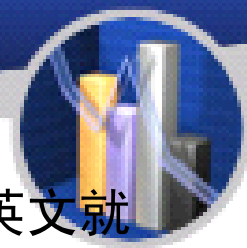
例如：巨、匠、国、连、原、问等。

五笔相关的名词解释



- **字根：** 汉字都是由笔划或部首组成的，一个汉字可以拆成一些最常用的基本单位，这个基本单位在五笔字型里就叫字根
- **字型：** 也就是构成汉字的字根之间的关系，在五笔字型编码方案里把字型分为三种结构，即上下型，左右型，杂合型。
- **成字字根：** 基本字根中有些字根本身就是一个完整的汉字，例如：日月人火手等，这种字根在五笔字型里叫做"成字字根"。
- **键名字根：** 在同一个键位上的基本字根中，选择一个具有代表性的字根，称为键名字根。五笔编码键盘图中各个键位左上角的字根。
- **键外字：** 不是成字字根或是键名字根的字，即是需要进行拆分的。绝大部分汉字都是键外字。

五笔相关的名词解释



- **字根码:** 每个字根都分派在一个字母键上，其所在键上的英文就是该字根的"字根码"。
- **简码:** 为了减少击键次数，提高输入速度而设置的。它将常用汉字只取其前边的一个、两个或三个字根构成，因为"识别码"总是在全码的最后位置，所以简码的设计不但减少了击键次数，而且省去了部分汉字的"识别码"的判别和编码，给击键带来了很大方便。
- **重码字:** 如果一个编码对应着几个汉字，这几个字称为重码字；
- **容错码:** 几个编码对应一个汉字，这几个编码称为汉字的容错码。

五笔字型的键盘设计



五笔字型的基本字根有**130**种，加上一些基本字根的变型，共有**200**个左右。

按照每个字根的起笔代号，分为**五个“区”**。

1区横区，**2区**竖区，**3区**撇区，**4区**捺区，**5区**折区。

每个区又分为**1、2、3、4、5**五个“位”，区和位对应的编号就称为“区位号”。

区位号用代码**11、12、13、14、15；21、22……51、52、53、54、55**表示

1区(横区): 王(11) 土(12) 大(13) 木(14) 工(15)

2区(竖区): 目(21) 日(22) 口(23) 田(24) 山(25)

3区(撇区): 禾(31) 白(32) 月(33) 人(34) 金(35)

4区(捺区): 言(41) 立(42) 水(43) 火(44) 之(45)

5区(折区): 己(51) 子(52) 女(53) 又(54) 纟(55)

五笔字型的键盘设计



| 第3区 (撇起笔) ← | | | | | 第4区 (点起笔) → | | | | |
|----------------|-----------|-----------|-----------|-----------|----------------|-----------|-----------|-----------|-----------|
| 金 35 Q | 人 34 W | 月 33 E | 白 32 R | 禾 31 T | 言 41 Y | 立 42 U | 水 43 I | 火 44 O | 之 45 P |
| 第1区 (横起笔) ← | | | | | 第2区 (竖起笔) → | | | | |
| 工 15 A | 木 14 S | 大 13 D | 土 12 F | 王 11 G | 目 21 H | 日 22 J | 口 23 K | 田 24 L | ⋮ ; |
| 第5区 (折起笔) ← | | | | | | | | | |
| Z | 纟 55 X | 又 54 C | 女 53 V | 子 52 B | 巳 51 N | 山 25 M | < , | > . | ? / |

金 𠂔 𠂔
𠂔 𠂔
儿 儿 夕 夕
35 Q

人 八
𠂔 𠂔
𠂔 𠂔
34 W

月 月
用 𠂔 乃 𠂔
豕 豕 𠂔 𠂔
33 E

白 𠂔
手 𠂔
斤 斤
32 R

禾 禾
竹 𠂔
𠂔 𠂔
31 T

言 𠂔
𠂔 𠂔 主
文 方 广
41 Y

立 𠂔
辛 𠂔 斗
六 𠂔 𠂔
42 U

水 𠂔
𠂔 𠂔
小 𠂔 𠂔 𠂔
43 I

火 𠂔
𠂔 𠂔 米
44 O

之 𠂔
𠂔 𠂔
𠂔 𠂔
45 P

工 𠂔
戈 𠂔 𠂔 七
𠂔 𠂔 𠂔 𠂔
15 A

木 丁
西
14 S

大 犬
手 𠂔 𠂔 古
𠂔 𠂔 𠂔 石
13 D

土 𠂔
𠂔 𠂔 𠂔
寸 𠂔 雨
12 F

王 𠂔
主 𠂔 五
戈 𠂔
11 G

目 𠂔
上 𠂔 𠂔 𠂔
止 𠂔 𠂔
21 H

日 𠂔
𠂔 𠂔 𠂔
早 𠂔 虫
22 J

口 𠂔
川
23 K

田 甲 𠂔
四 𠂔 𠂔 𠂔 𠂔
𠂔 𠂔 𠂔 𠂔 𠂔
24 L

学习键
66 Z

𠂔 𠂔 𠂔 𠂔
𠂔 𠂔
𠂔 𠂔
55 X

又 𠂔 𠂔 𠂔
巴 𠂔 𠂔
54 C

女 𠂔
刀 𠂔 𠂔
九 𠂔 𠂔
53 V

子 𠂔
耳 𠂔 𠂔 𠂔 𠂔
了 𠂔 𠂔
52 B

已 𠂔 𠂔 𠂔
尸 𠂔 𠂔 𠂔
心 𠂔 𠂔 𠂔
51 N

山 𠂔
𠂔 𠂔
𠂔 𠂔
25 M

35 ← 31
3区(撇起笔)

41 → 45
4区(点起笔)

15 ← 11
1区(横起笔)

21 → 24
2区(竖起笔)

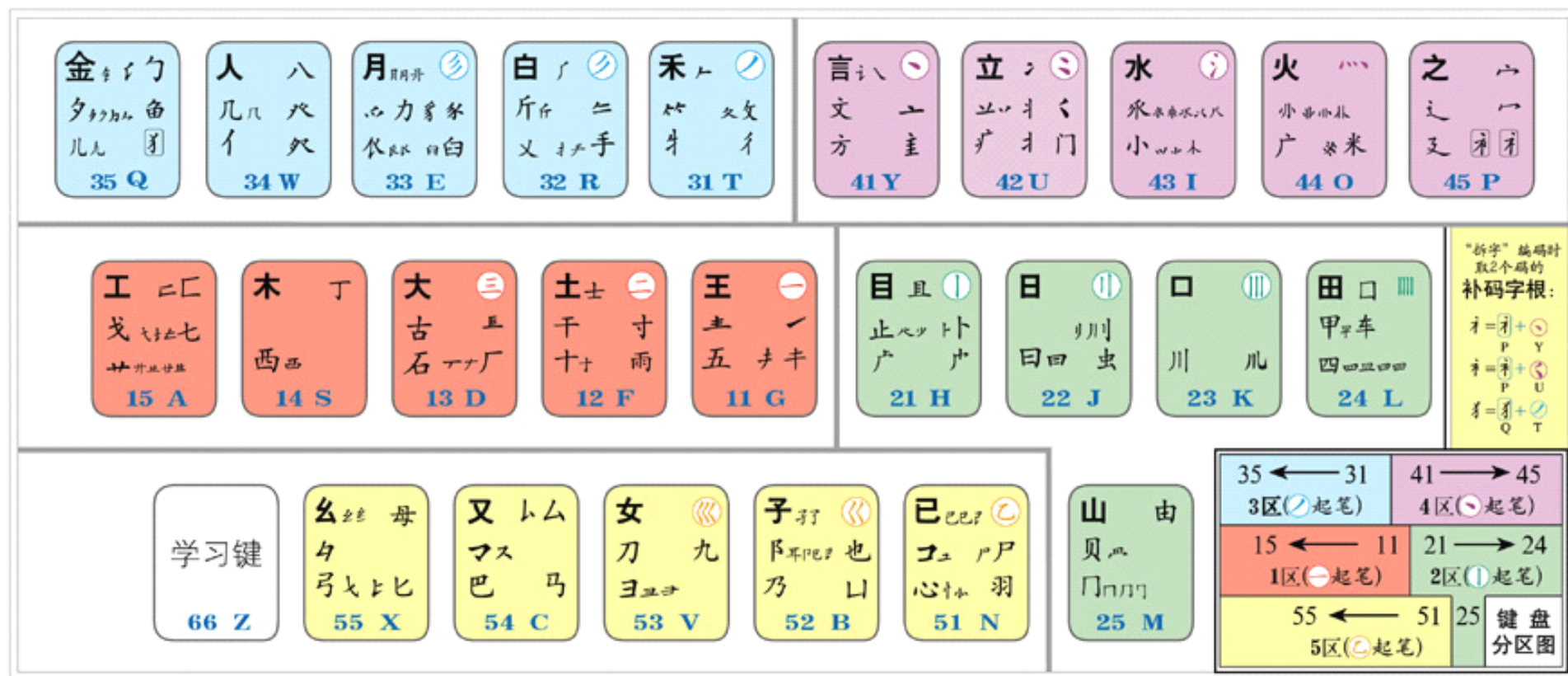
55 ← 51
5区(折起笔)

25 键
盘
分区图

86版

| 1(横)区 | 2(竖)区 | 3(撇)区 | 4(捺)区 | 5(折)区 |
|---|---|---|--|---|
| 11G 王旁青头戈五一 12F 土土二千十寸雨 13D 大犬三羊古石厂 14S 木丁西 15A 工戈草头右框七 | 21H 目具上止卜虎皮 22J 日早两竖与虫依 23K 口与川，字根稀 24L 田甲方框四车力 25M 山由贝，下框几 | 31T 禾竹一撇双人立 反文条头共三一 32R 白手看头三二斤 33E 月多乃用家衣底 34W 人和八，三四里 35Q 金勺缺点无尾鱼 犬旁留义儿一点夕 氏无七 | 41Y 言文方广在四一 高头一捺谁人去 42U 立辛两点六门广 43I 水旁兴头小倒立 44O 火业头，四点米 45P 之宝盖，摘衤（衣） | 51N 已半已满不出已 左框折尸心和羽 52B 子耳了也框向上 53V 女刀九白山朝西 54C 又巴马，丢矢矣 55X 慈母无心弓和匕 幼无力 |

第三代 五笔字型® (新世纪版) 简体字根键位图



1区: ㇀起笔

2区: ㇀起笔

3区: ㇀起笔

4区: ㇀起笔

5区: ㇀起笔

11 G 王旁青头五一提

21 H 目止具头卜虎皮

31 T 禾竹牛旁卧人立

41 Y 言文方点在四一

51 N 己类左框心尸羽

12 F 土士二千十寸雨

22 J 日曰两竖与虫依

32 R 白斤气头叉手提

42 U 立带两点病门里

52 B 子耳了也乃齿底

13 D 大三肆头古石厂

23 K 口中两川三个竖

33 E 月舟衣力豕豸白

43 I 水边一族三点小

53 V 女刀九巡录无水

14 S 木丁西边要无女

24 L 田框四车甲单底

34 W 人八登祭凤头几

44 O 火变三态广二米

54 C 又巴甬矣马失蹄

15 A 工戈草头右框七

25 M 山由贝骨下框里

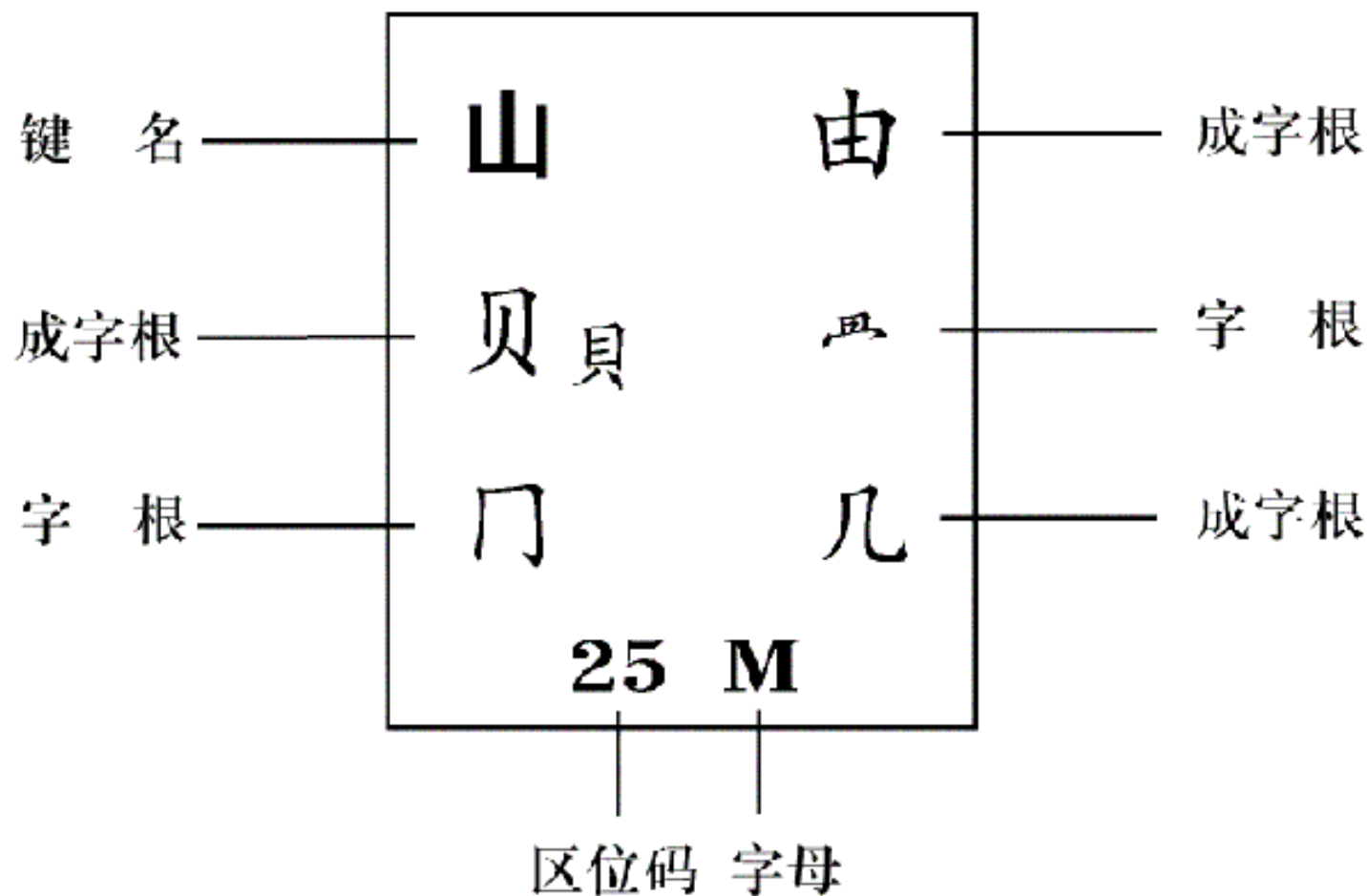
35 Q 金夕才儿包头鱼

45 P 之字宝盖补示衣

55 X 么母绞丝弓三匕

声明: 以上字根键位图及助记歌, 王码公司拥有专利、版权和著作权, 未经许可不得为商业目的转载和发行 2000年5月6日。

五笔字型字根分布的助记特征



五笔字型字根分布的助记特征



(1)字根的首笔代号与所在的区号一致。

如：

| | | | |
|---|---|---|---|
| 王 | 在 | 1 | 区 |
| 木 | 在 | 1 | 区 |
| 目 | 在 | 2 | 区 |
| 金 | 在 | 3 | 区 |
| 之 | 在 | 4 | 区 |
| 乚 | 在 | 5 | 区 |

五笔字型字根分布的助记特征



(2)次笔代号与所在的位号一致

如：

| | | | | | |
|---|---|---|---|---|---|
| 王 | 在 | 1 | 区 | 1 | 位 |
| 土 | 在 | 1 | 区 | 2 | 位 |
| 石 | 在 | 1 | 区 | 3 | 位 |
| 之 | 在 | 4 | 区 | 5 | 位 |
| 广 | 在 | 4 | 区 | 1 | 位 |
| 么 | 在 | 5 | 区 | 5 | 位 |

五笔字型字根分布的助记特征



(3)基本字根的笔画数与位号一致

一、丨、丿、㇏、乙，
这5个单笔画，都在第1位；

二、㇀、㇁、㇂、㇃、㇄，
这5个双笔画，都在第2位；

三、川、彡、彣、彤、
这5个三笔画，由3个单笔构成，都在第3位；

川丨、灬、
这5个四笔画，由4个单笔构成，都在第4位。

| | 1位 | 2位 | 3位 | 4位 |
|----|----|----|----|----|
| 1区 | 一 | 二 | 三 | |
| 2区 | 丨 | ㇀ | 川 | 川丨 |
| 3区 | 丿 | ㇂ | 彡 | |
| 4区 | ㇏ | ㇄ | 彣 | 灬 |
| 5区 | 乙 | ㇄ | 彣 | |

五笔字型字根分布的助记特征



(4)基本字根与键名字根形态相似或相近

如："王"字键上，有"五，戈"等字根，"日"字键上有"虫、早"等字根。

王
土
田
之
大
月
已

与
与
与
与
与
与
与

五
士
甲、
乚、
犬
月、
己、

四
彳
用
巳、
尸

五笔字型字根分布的助记特征



(5)个别例外

- 带有两点立 六 辛 丩 等字根都在 4 区 2 位上
- 车，力在L键上（繁体车与甲相似，力的声母为L）
- 心在N键上，心字最长的笔画为折笔

五笔字型字根分布的助记特征



| | | | | |
|--|---|---|---|---|
| 工 匚 戈 𠂇 𠂉 𠂊 𠂋 𠂌 𠂍 𠂎 𠂏 15 A | 木 丁 西 14 S | 大 犬 ③ 𠂐 𠂑 𠂒 𠂓 𠂔 𠂕 𠂖 𠂗 𠂘 𠂙 𠂚 13 D | 土 士 ② 𠂛 𠂜 𠂝 𠂞 𠂟 𠂠 𠂡 𠂢 𠂣 12 F | 王 ① 𠂤 𠂥 𠂦 𠂧 𠂨 𠂩 𠂪 𠂫 11 G |
|--|---|---|---|---|

| 分区 | 区位 | 键位 | 识别码 | 标识字根 | 键名↓ | 字 根 | 助 记 词 | 高频字 |
|-----------------------|----|----|-----|------|-----|---|---------------|-----------------------|
| 一 区 横 起 笔 | 11 | G | ① | 一 / | 王 | 主 戈 五 一 | 王旁青头 戈 五 一 | 一 地 在 要 工 |
| | 12 | F | ② | 二 | 土 | 士 干 十 半 帛 寸 雨 | 土 士 二 干 十 寸 雨 | |
| | 13 | D | ③ | 三 𠂔 | 大 | 犬 𠂐 𠂑 𠂒 𠂓 𠂔 𠂕 𠂖 𠂗 𠂘 𠂙 𠂚 | 大 犬 三 羊 古 石 厂 | |
| | 14 | S | | | 木 | 丁 西 西 | 木 丁 西 在 一 四 里 | |
| | 15 | A | | | 工 | 戈 𠂇 𠂉 𠂊 𠂋 𠂌 𠂍 𠂎 𠂏 𠂐 𠂑 𠂒 𠂓 𠂔 𠂕 𠂖 𠂗 𠂘 𠂙 𠂚 | 工 戈 草 头 右 框 七 | |

注： 1. 助记词中，11.G键上的“戈”读“兼”，13.D键上“羊”指“平”。

五笔字型字根分布的助记特征



| | | | | |
|---------------------------------|---------------------------------|------------------|---|-----------------------------|
| 目 且 ① 上 卜 尸 止 乚 尸 21 H | 日 曰 四 ② 川 川 早 月 虫 22 J | 口 ③ 川 23 K | 田 甲 ④ 四 四 四 四 四 四 口 车 车 力 24 L | 山 由 贝 贝 几 门 几 25 M |
|---------------------------------|---------------------------------|------------------|---|-----------------------------|

| 分区 | 区位 | 键位 | 识别码 | 标识字根 | 键名↓ 字 根 | 助 记 词 | 高频字 |
|-------------------|----|----|-----|---------|-------------------------|---------------|-----|
| 二区 竖 起 笔 | 21 | H | ① | 丨 ↓ | 目 止 乚 且 上 卜 尸 尸 | 目 止 具 头 卜 虎 皮 | 上 |
| | 22 | J | ② | ㇀ ㇀ ㇀ ㇀ | 日 早 曰 四 ㇀ ㇀ ㇀ ㇀ ㇀ ㇀ | 日 早 两 竖 与 虫 依 | 是 |
| | 23 | K | ③ | ㇀ ㇀ | 口 川 | 口 中 一 川 三 个 竖 | 中 |
| | 24 | L | | ㇀ | 田 甲 口 四 四 四 四 四 四 车 车 力 | 田 甲 方 框 四 车 力 | 国 |
| | 25 | M | | | 山 由 几 贝 贝 门 几 | 山 由 贝 骨 下 框 几 | 同 |

五笔字型字根分布的助记特征



| | | | | |
|---|---|---|---|---|
| 金 钅 钅 乚 乚 儿 儿 夕 夕 35 Q | 人 八 人 人 亻 亻 34 W | 月 月 夕 用 舟 乃 心 冫 冫 冫 冫 冫 冫 33 E | 白 厂 夕 手 扌 扌 斤 斤 32 R | 禾 禾 夕 竹 竹 人 文 文 亻 31 T |
|---|---|---|---|---|

| 分区 | 区位 | 键位 | 识别码 | 标 识 字 根 | 键名 ↓ 字 根 | 助 记 词 | 高 频 字 |
|-------------------|----|----|-----|------------|-----------------------|---------------------|-----------------------|
| 三区 撒 起 笔 | 31 | T | ㄣ | ノ ㄣ | 和 禾 竹 人 文 文 亻 | 禾 竹 反 文 双 人 立 | 和 的 有 人 我 |
| | 32 | R | ㄥ | ㄥ 厂 ㄣ | 白 厂 斤 斤 ㄣ 手 扌 扌 | 白 斤 气 头 手 边 提 | |
| | 33 | E | ㄣ | ㄣ 心 | 月 乃 用 舟 冫 冫 冫 冫 冫 冫 | 月 乃 用 舟 冫 冫 冫 冫 冫 冫 | |
| | 34 | W | | | 人 亻 八 人 人 | 人 八 登 祭 把 头 取 | |
| | 35 | Q | | | 金 钅 夕 夕 乚 乚 儿 儿 乚 乚 钅 | 金 夕 乚 儿 包 头 鱼 | |

注： 1. 助记词中，35.Q键上的“乚”读“叉”

五笔字型字根分布的助记特征



言 讠言 ㄣ
 讠 讠 主
 文 ㄣ 方 广
41 Y

立 立 ㄣ
 辛 冫 斗
 六 冫 冫 冫
42 U

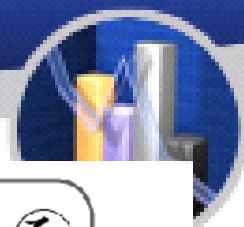
水 水 ㄣ
 冫 冫 冫
 小 冫 冫 冫
43 I

火 ㄣ
 业 冫 冫 米
44 O

之 ㄣ
 之 冫 冫
 之 冫 冫
45 P

| 分区 | 区位 | 键位 | 识别码 | 标识字根 | 键名↓ 字 根 | 助 记 词 | 高频字 |
|-------|----|----|-----|-------|--|---------------|-----------|
| 四区点起笔 | 41 | Y | ㄣ | 、 \ | 言 <small>讠言</small> 讠 讠 主 文 ㄣ 方 广 | 言 文 方 广 在 四 一 | 主 产 不 为 这 |
| | 42 | U | ㄣ | 、 冫 冫 | 立 <small>立</small> 辛 冫 斗 六 冫 冫 冫 冫 冫 冫 冫 冫 | 立 辛 两 点 病 门 里 | |
| | 43 | I | ㄣ | 冫 冫 | 水 <small>水</small> 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 | 水 族 三 点 兴 头 小 | |
| | 44 | O | | ㄣ | 火 <small>火</small> 业 冫 冫 米 | 火 里 业 头 四 点 米 | |
| | 45 | P | | | 之 <small>之</small> 之 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 冫 | 之 字 宝 盖 补 冫 冫 | |

五笔字型字根分布的助记特征



纟 纟 纟 纟 纟
 纟 纟 纟 纟 纟
 弓 纟 比
 55 X

又 ス ム
 巴 馬 馬
 54 C

女 ㄣ
 刀 白 白
 九 ヨ ヨ
 53 V

子 子 ㄣ
 耳 阝 也
 了 囟 口
 52 B

己 己 己 乙
 尸 尸 ㄣ
 心 忄 羽
 51 N

| 分区 | 区位 | 键位 | 识别码 | 标识字根 | 键名↓ 字 根 | 助 记 词 | 高频字 |
|-------|----|----|-----|------|---------------------|---------|-----|
| 五区折起笔 | 51 | N | 乙 | 乙 | 己 己 己 ㄣ 心 忄 尸 尸 羽 | 己类左框心尸羽 | 民 |
| | 52 | B | ㄣ | ㄣ | 子 子 耳 阝 了 也 口 囟 | 子耳了也框上举 | 了 |
| | 53 | V | ㄣ | ㄣ | 女 刀 九 白 白 ヨ ヨ | 女刀九ㄣ白山倒 | 发 |
| | 54 | C | | | 又 巴 ス ム 馬 馬 | 又巴劲头私马依 | 以 |
| | 55 | X | | | 纟 纟 纟 纟 纟 弓 比 比 纟 纟 | 绞丝互腰弓和匕 | 经 |

注：1. 助记词中，53.V键上的“白”读“旧”，54.C键上“私”指“ム”，55.X键上的“纟”读“腰”、“匕”读“互腰”。

五笔字型汉字拆分



汉字与字根的组成关系

- (1) 汉字结构层次
(笔画、字根、汉字)
- (2) 基本笔画
(横、竖、撇、捺、折)
- (3) 基本字型
(左右型、上下型、杂合型)

由两个字根组成的汉字称作双合字，
由三个字根组成的汉字称作三合字，
由四个字根组成的汉字称作四合字，
由多于四个字根组成的汉字称作多合字。

五笔字型汉字拆分



1. 五笔字型基本笔画

在五笔字型编码方案中，汉字的笔划只归结为横、竖、撇、捺、折五种。

因为"点"和"捺"两者运笔方向基本一致，所以把"点"归为"捺"；把挑（提）归结于"横"类；左钩归为"竖"，其他带转折的笔划都归结为"折"类。

| 笔划名称 | 笔划代码 | 笔划走势 | 笔划及其变形 |
|------|------|-------|----------|
| 横 | 1 | 左→右 | 横一提/ |
| 竖 | 2 | 上→下 | 竖丨 竖左勾 |
| 撇 | 3 | 右上→左下 | 撇 |
| 捺 | 4 | 左上→右下 | 捺 点 |
| 折 | 5 | 带转折 | 各种带转折的笔划 |

五笔字型汉字拆分



2. 汉字的字型信息

1型： 左右部位结构的汉字，

例如：明、响、构等。

虽然"构"的右边"勾"是两个基本字根按内外型组合成的，但整字仍属于左右型，这种分类主要是根据汉字的偏旁部首位置来进行的。

2型： 部位结构是上下型的字，

例如：昌、感、菱、巍等。

同左右型一样，像"巍"这类字虽然下面的"魏"是按左右型组合成的，但整字仍是上下型。

3型： 称为杂合型。

包括单体结构和内外结构的汉字，也就是没有明显上下和左右结构特点的汉字。

五笔字型汉字拆分



左右型：

- (1) 在由两个字根组成的汉字，即双合字中，两个字根分列左右，整个汉字中有着明显的界线，并且字根间有一定的距离，
如：汉、明、林、极、休等。
- (2) 在三合字中，组成整个汉字的三个字根从左到右排列，或者单独占据一边的一个字根与另外两个字根呈左右排列，
如：辩、掰等。
- (3) 在四合字或多合字中，组成整个汉字的若干字根很明显地分成左右两部分，无论左右那一边字根数多，都将这种汉字定为左右类结构的汉字，
如：械、讹、键等。

五笔字型汉字拆分



上下型：

- (1) 在双合字中，两个字根的位置是上下的关系，这两个字根之间有着明显的界线，且有一定的距离，
如：节、个、字、另等。
- (2) 在三合字中，三个字根也是分成两个部分，虽然其中一个部分的字根数要多一些，但两个部分仍然是上下两层的位置关系，
如：意、花、怒、想等。
- (3) 组成四合字或多合字的字根也明显地分成上下两部分，则无论是上半部分字根数多一些或是或是下半部分字根数多一些，这样的汉字也属于上下结构型汉字，
如：赢、离、聚等。

五笔字型汉字拆分



杂合型：

杂合型汉字包括**单体**，**内体**，**包围**三种类型，指组成整字的各个字根之间没有简单明确的左右或上下型关系，

如：因、周、半、太等，象这样的汉字，字根之间都是内外或包围的关系。

在五笔字型汉字结构的划分中，必须着重注意下面几个约定：

- (1) 凡单笔画与字根相连者或带点结构都视为杂合型。
- (2) 汉字结构区分时，也要按"能散不连"的原则来进行。如"矢、卡、严"都视为上下型。
- (3) 含两字根且相交者属杂合型，如"乐、串、电、本、无、农"。
- (4) 下含"走之"字为杂合型，如"进、过，遂"等。
- (5) 以下各字为杂合型：司、床、厅、龙、尼、式、后、处等，但相似的右、左、有、布、灰等可视上下型。

五笔字型汉字拆分



3.字根间的位置关系

(1)单 基本字根就单独构成一个汉字。

如：“金、王、口、皿、木、火、目”。

(2)散 构成汉字不止一个字根，且字根之间保持一定的距离，不相连也不相交。

如：“讲、肥、明、张、吴”等。

(3)连 一个基本字根连一单笔画，或者是“带点结构”

如：“自”，“千”，“于”，“本”等。

如：“勺”、“术”、“太”、“主”等。

(4)交 多个字根交叉迭构成汉字。

如：“申”，“里”，“丰”，“必”，“专”等。

五笔字型汉字拆分



4.五笔字型汉字拆分原则

拆字依据，书写顺序，
左右上下，分别拆取。
取大优先，兼顾直观，
能连不交，能散不连。

五笔字型汉字拆分



(1) “书写顺序”

谁 → 讠 亻 一 (从左到右, 左右型)
茶 → 艹 人 木 (从上到下, 上下型)
问 → 门 口 (从外到内, 杂合型)
夷 → 一 弓 人 (书写顺序)

(2) “取大优先”

果 → 曰 木 (正确拆分)
 → 曰 十 八 (错误拆分)

(3) “兼顾直观”

自 → 丿 目
国 → 口、王、丶

(4) “能连不交”

天 → 一 大 (两个字根互相连接, 正确拆分)
天 → 二 人 (两个字根互相交叉, 错误拆分)

(5) “能散不连”

占 → 卜 口 (两个字根关系是散开的, 都不是单笔画)
非 → 三 丨 三 (三个字根关系是散开的, 都不是单笔画)

五笔字型的汉字编码



五笔字型单字编码歌诀

五笔字型均直观，依照笔顺把码编；
键名汉字打四下，基本字根请照搬；
一二三末取四码，顺序拆分大优先；
不足四码要注意，交叉识别补后边。

五笔字型的汉字编码



五笔字型拆字取码的五项原则：

- 1、从形取其顺序按书写规则，即从左到右、从上到下、从外到内；
- 2、以**130**多个字根为基本单位；
- 3、字根数超过四个的汉字，按一二三末字根的顺序，最多只取四码；
- 4、单体结构拆分取大优先；
- 5、末笔与字型交叉识别。

五笔字型汉字编码流程图

汉字

键面有

键名：把所在键连打四下

例：言：言言言言；又：又又又又；山：山山山山

YYYY CCCC MMMM

成字根：报户口 加打第一、二、末单笔画（不足四码，补空格键）

例：西：西一丨一；方：方丶一乙；厂：厂一丨（空格）

SGHG YYGN DGT

键面无

依照

书写顺序

例：

新：立木斤（正）

立斤木（误）

取大优先

例：

平：一丨丨（正）

一丨丨（误）

兼顾直观

例：

自：丨目（正）

丨目（误）

能连不交

例：

天：一大（正）

二人（误）

拆

超过四码：取1、2、3、末字根编码

例：赣：立早欠贝；攀：木义手

刚好四码：依次键入即可

例：照：日刀口；到：一厶土丨

不足四码：字根键入完后，补打“末笔字型识别码”（仍不足四码，补空格键）

例：

汉：43 54 41

彳 又 丶

I C Y

会：34 12 54 42

人 二 厶 丨

W F C U

必：51 31 33

心 丨 彡

N T E

末笔字型识别码表

| 末笔形 | 字型 | 左右型 | 上下型 | 杂合型 |
|-----|----|--------|---------|----------|
| | | 1 | 2 | 3 |
| 横 | 1 | 11 (一) | 12 (二) | 13 (三) |
| 竖 | 2 | 21 (丨) | 22 (丨丨) | 23 (丨丨丨) |
| 撇 | 3 | 31 (丿) | 32 (丿) | 33 (彡) |
| 捺 | 4 | 41 (㇏) | 42 (㇏) | 43 (㇏) |
| 折 | 5 | 51 (乙) | 52 (乚) | 53 (乚) |

五笔字型的汉字编码



键名汉字的编码规则

五笔字型中规定的键名汉字共有**25**个，

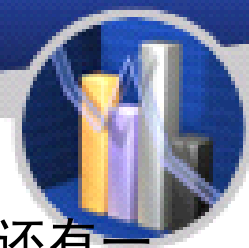
"王土大木工目日口田山禾白月人金言立水火之已子女乚纟"

25个键名汉字与**25**个字母键相应，这些字的编码相当简单，它们的编码就是**4**个所在字母键字母，

如："言"字的编码为"**YYYY**"，

"纟"字的编码为"**XXXX**"。

五笔字型的汉字编码



成字字根的编码规则

在五笔字型字根键盘的每个字母键上，除了一个键名字根外，还有一些其它类型的字根。有些字根其本身就是一个汉字，这样的字根称为成字字根。

成字字根的编码是按下面公式进行的：

键名码+首笔码+次笔码+末笔码

当成字字根仅为两笔时，编码只有三码，公式为：

键名码+首笔码+末笔码

例如："石"，字根所在的键为**D**，第一笔划为"横"的编码是**G**，第二笔为"撇"的编码是**T**键，末笔"横"的编码也是**G**键，所以这个字的五笔编码就是**DGTG**。

再比如"虫"字，这个字根是在**J**键，第一笔是竖**H**，第二笔是折**N**，最末一笔是捺(在五笔里"点"做为"捺")**Y**，所以"虫"的五笔编码就是**JHNY**。

五笔字型的汉字编码



五种单笔画的编码规则

一 : 11 11 24 24 (GGLL)

丨 : 21 21 24 24 (HHLL)

丿 : 31 31 24 24 (TTLL)

丶 : 41 41 24 24 (YYLL)

乙 : 51 51 24 24 (NNLL)

五笔字型的汉字编码



键外字的编码规则

五笔字型汉字编码主要是键外字的编码，编码可以分为两类，纯字根码和识别码。

(1) 纯字根码

如果一个汉字的字根是四个或超过四个，就用前三后一总共四个字根码组成编码。

(2) 字型结构识别码

不足四个字根的汉字需补一个字型结构识别码，以增加区分汉字的信息量。

五笔字型的汉字编码



纯字根码

每个字根都分派在一个字母键上，其所在键上的英文就是该字根的"字根码"。

凡含四个或超过四个字根的汉字，取其第一、二、三、末四个字根码组成键外字的输入编码。

第一字根编码+第二字根编码+第三字根编码+最末一个字根编码

这里一、二、三、末应按正常书写顺序，先左后右，先上后下，先外后内。

五笔字型的汉字编码



交叉识别码

一个字的字根不足四个时，依次击入字根码后，最后需要补一个识别码。

| | 左右结构 (第1位) | 上下结构 (第2位) | 杂合结构 (第3位) |
|------------|---------------|---------------|---------------|
| 末笔是“横”，在1区 | 11 G | 12 F | 13 D |
| 末笔是“竖”，在2区 | 21 H | 22 J | 23 K |
| 末笔是“撇”，在3区 | 31 T | 32 R | 33 E |
| 末笔是“捺”，在4区 | 41 Y | 42 U | 43 I |
| 末笔是“折”，在5区 | 51 N | 52 B | 53 V |

五笔字型的汉字编码



1. 丢失字型信息会引起重码

叭： 口 八 (KW) (1型字)

只： 口 八 (KW) (2型字)

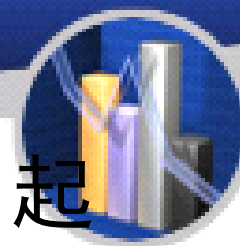
吧： 口 巴 (KC) (1型字)

邑： 口 巴 (KC) (2型字)

旡： 九 日 (VJ) (2型字)

旭： 九 日 (VJ) (3型字)

五笔字型的汉字编码



2. 因同一个键位上有几个不同的码元而引起重码

沐： 氵 木 (IS) (末笔为 丶)

汀： 氵 丁 (IS) (末笔为 丨)

洒： 氵 西 (IS) (末笔为一)

他： 亻 也 (WB) (末笔为 乙 , 1型字)

仓： 人 卩 (WB) (末笔为 乙 , 2型字)

仔： 亻 子 (WB) (末笔为一, 1型字)

五笔字型的汉字编码



末笔画识别码的规定

(1) 对于由“乚”、“乚”、“门”、“犭”组成的半包围汉字，以及由“口”组成的全包围汉字，末笔一律规定为被包围部分笔画结构的末笔。

如：“进”、“逞”、“选”、“建”、“国”等字。

(2) 末字根为“力”、“刀”、“九”、“匕”等时，一律认为末笔画为折。

如：花 **AWXB**

仇 **WVN**

边 **LPV**

(3) “我”、“戈”，“成”等字的末笔取“丿”。

如：我 **TRNT** 末笔为撇

五笔字型的汉字编码



一级简码

五笔字型中，根据每个字母键上的字根形态特征，每键安排一个最为常用的高频汉字，这类字共**25**个，它们的编码只有一位，输入时只要击键一次再加一次空格键即可。这些高频字及其编码如下：

第一区： 一 -> 王 地 -> 土 在 -> 大 要 -> 木 工 -> 工
第二区： 上 -> 目 是 -> 日 中 -> 口 国 -> 田 同 -> 山
第三区： 和 -> 禾 的 -> 白 有 -> 月 人 -> 人 我 -> 金
第四区： 主 -> 言 产 -> 立 不 -> 水 为 -> 火 这 -> 之
第五区： 民 -> 已 了 -> 子 发 -> 女 以 -> 又 经 -> 纟

五笔字型的汉字编码



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

五笔字型词组编码



五笔字型词组编码规则

(1) 二字词组的编码

二字词的编码由所含的两个汉字各取前两个字根码组成，即每个汉字按书写顺序取前两个字根来进行编码。

如：

机器：木 几 口 口 **SMKK**

计算：言 十 竹 目 **YFTH**

数量：米 女 日 一 **OVJG**

(2) 三字词的编码

三字词的编码由前两个汉字的第一个字根码和后一个汉字的前两个字根码构成。

如：

计算机：言 竹 木 几 **YTSM**

工艺品：工 艹 口 口 **AAKK**

五笔字型词组编码



(3) 四字词的编码

四字词的编码是由每个汉字的前一字根码组成，共四码，

例如：

巧夺天工：工 大 一 工 **ADGA**

原原本本：厂 厂 木 木 **DDSS**

(4) 多字词的编码

多字词是指构成词的单个汉字数超四个，多个词的编码按“一、二、三、末”的规则，即分别取第一、第二、第三及最末一个汉字的第一个字根码来构成的编码。

例如：

中华人民共和国：口 亻 人 口 **KWWL**