

ISO/IEC JTC1/SC2/WG2

N937

TITLE: A Proposal For Adding Combining Symbols
To Represent the Han Characters

TO: ISO/IEC JTC1/SC2/WG2

SOURCE: CHINA

REQUESTED ACTION: CONSIDERATION BY SC2/WG2

BEIJING

24 October 1993

A Proposal For Adding Combining Symbols in BMP To Represent the Han Characters

1. OBJECTIVES

Han characters are usually classified as an ideographically based language compared to alphabetically based languages, such as English and French. However, Han characters can be constructed by a series of symbols that describes both the Han roots that form a character and the relative positions of these Han roots. Thus Han characters can be considered as a two dimensional alphabetically based language. The construction of new Han characters does follow certain rules and symbols to describe such construction is finite, very much like the construction of new English words. Study show that the number of symbols that are used to describe Han characters does not go beyond 2K. Based on these, we propose to put combining symbols, including both the Han roots and the structure symbols to the BMP in UCS so that all Han characters beyond the BMP can be presented. As the current BMP already included a part of Han character Han roots, we only request an additional 2K code points to put the combining symbols.

The rest of the proposal is organized as the following.

Section 2 describes the theoretical basis of Sequential Description of Han character Method. Practices using this method are also given in Section 2. Section 3 gives the justification for using this method and comparison to other languages in BMP. Section 4 is a conclusion.

2. THE COMBINING SYMBOL METHOD AND ITS PRACTICE

Although Han characters are classified as ideographically based language, only Han roots are considered as indivisible symbols. Only a subset of these Han roots can be used alone as Han characters. In general, the majority of Han characters must use more than one Han roots and the arrangement of these Han roots is not one dimensional like English, the relative position of these Han roots in a two dimensional area must also be specified. The following shows a Chinese antithetical phrase:

山石岩下古木枯,此木是柴
白水泉边女子好,少女最妙

For any person who understand Han characters, this phrase shows that characters can be formed from Han roots. These Han roots can also be used as characters themselves. For instance,

岩=山+石, 泉=白+水
枯=木+古, 好=女+子
柴=此+木, 妙=女+少

In the previous example, the relative positions of different Han roots do not need to be described given as the intrinsic property of two dimensional expression clearly indicates the relative positions of the Han roots. In order to linearize the description of Han characters for modern communication and processing, some methods must be developed. Section 2.1. formally introduces such a method. Section 2.2 describes how this method is already used in practice in the communities where Han characters are used.

2.1 The Method

From studies results of Han character constructions, we have concluded that the Han root symbols and structure symbols used Han character does not go beyond 2,000.

Before we go further with the discussion we first induce formal definitions of some terms used in this proposal.

Definition 1: Combining Graphic Symbol

A combining Graphic Symbol is a combination of two or more graphic symbols at one Han character position.

For example: "吳" is a combination of the two graphic symbols "口" and "天".

Definition 2: Han Root

Han roots are smallest building elements, that are drawn from Han characters according to the patterns of their strokes.

For example: "月" and "生" are two Han roots of Han character "胜".

Definition 3: Han Root Symbol

A Han root Symbol is the visual representation of a Han root.

Definition 4: Han Root Block

A Han root block is a combination of one or more Han root symbols in a plane, according to their relative positions.

Definition 5: Han Character Structure

Han character structure is used to describe the relations between Han root blocks and their relative positions. Including the following building elements:

Definition 5.1: Single Structure

If Han character has only one Han root symbol, its structure is called single structure.

For example: The Han root "王" is also the Han character "王".

Definition 5.2: Left-right Structure

If a Han character has two Han root blocks, and one of the two Han root blocks is on the left and the other is on the right, its structure is called Left-right structure.

For example: Two Han roots "口" and "巴" can be used by the Left-right structure to form the Han character "吧".

Definition 5.3: Upper-lower Structure

If a Han character has two Han root blocks, and one of the two Han root blocks is on the upper position, while the other is on the lower position, its structure is called Upper-lower structure.

For example: The Han roots "口" and "巴" can be used by the Upper-lower structure to form the Han character "邑" which means completely different this than the Han character "吧" given the previous example.

Definition 5.4: Enclosing Structure

If a Han root block encloses another Han root block by two or more sides, the structure is called Enclosing structure. Following are seven structures:

Definition 5.4.1: Surround structure

If a Han root block encloses another Han root block by four sides, this structure is called surround structure.

For example: Two Han roots "口" and "玉" can be used by the Surround structure to form the Han character "国".

Definition 5.4.2: Top-enclosing Structure

If a Han root block encloses another Han root block by left, top and right sides, the structure is called Top-enclosing structure.

For example: Two Han roots "冂" and "凵" can be used by the Top-enclosing structure to form the Han character "冈".

Definition 5.4.3: Bottom-enclosing Structure

If a Han root block encloses another Han root block by left, bottom and right sides, the structure is called Bottom-enclosing structure.

For example: Two Han roots "凵" and "凵" can be used by the Bottom-enclosing structure to form the Han character "凶".

Definition 5.4.4: Left-enclosing Structure

If a Han root block encloses another Han root block by top, left and bottom sides, the structure is called Left-enclosing structure.

For example: Two Han roots "匚" and "斤" can be used by the Left-enclosing structure to form the Han character "匠".

Definition 5.4.5: Top-left-enclosing Structure

If a Han root block encloses another Han root block by top and left sides, the structure is called Top-left-enclosing structure.

For example: Two Han roots "广" and "车" can be used by the Top-left-enclosing structure to form the Han character "库".

Definition 5.4.6: Top-right-enclosing Structure

If a Han root block encloses another Han root block by top and right sides, the structure is called Top-right-enclosing structure.

For example: Two Han roots "冂" and "口" can be used by the Top-right-enclosing structure to form the Han character "句".

Definition 5.4.7: Bottom-left-enclosing Structure

If a Han root block encloses another Han root block by left and bottom sides, the structure is called Bottom-left-enclosing structure.

For example: Two Han roots "走" and "干" can be used by the Bottom-left-enclosing structure to form the Han character "赶".

Definition 5.5: Pile Structure

If two Han root blocks pile up, the structure is called Pile structure. This structure is used to represent some special Han characters.

For example: Two Han roots "口" and "十" pile up, special symbol "𠔁" is combined.

Definition 6: Han Character Structure Symbol

A Han character structure symbol is the visual representation of a Han character structure, called structure symbol.

For example:

Upper-lower structure symbol is □

Left-right structure symbol is □

Surroud structure symbol is □

Note:A structure symbol is a graphic character.

Definition 7: Stroke

A stroke is an element used to form the shape of Han characters
.It is a smallest element of a Han root.

Definition 8: Sequence

Sequences refer to:

Definition 8.1: The sequence of strokes: — |) \ 乙

Definition 8.2: The sequence of Han root codes: fewer strokes first, more strokes second, if the number of strokes is the same, follow the sequence of strokes.

Definition 8.3: The Sequence of Han Root Blocks

- . The sequence in upper-lower structure: upper first, lower second.
- . The sequence in left-right structure: left first, right second.
- . The sequence in enclosing structure: out first, inner second.
- . The sequence in pile structure: same as the sequence of Han root codes.

Definition 9: Sequential Han Root String (called Han Root String)

A code combination, composed of structure symbol and Han root symbol according to their sequences, is called a Han root string.

The representation of Han characters using sequential Han root combination method:

One Han character can be represented by using a sequential Han root string. This can be described as:

represent

<a Han character> ——> <sequential Han root string>

The sequential Han root string is composed of structure symbols and Han roots symbols.

There are three forms:

Form 1: To represent single structure Han character.

The sequential Han root string, used to represent single structure Han character, is composed of a structure symbol and a Han root symbol. In this sequential Han root string, the first symbol is a structure symbol, the second is a Han root symbol. But the structure symbol can be omitted. This can be described as:

represent

<Sequential Han root string> ——> <structure symbol>
<Han root symbol>

For example: Han character Sequential Han root string

丈 □丈

 or 丈

Form 2: To represent a Han character which uses only one structure symbol.

The sequential Han roots string,used to represent a Han character which uses only one structure symbol,is composed of a structure symbol and two Han root symbols.The first one must be a structure symbol.The sequence of the two Han root symbols is the same as that of Han root blocks.This can be described as:

represent

<Sequential Han root string> →
 <structure symbol>
 <Han root symbol>
 <Han root symbol>

For example: Han character Sequential Han root string

好	☐女子
岩	☐山石
国	☐口玉

Form 3: To represent a Han character which uses several structure symbols

The sequential Han roots string,used to represent a Han character which uses several structure symbols,is composed of a structure symbol and two Han root strings. The first symbol must be a structure symbol,and any one of the two Han root strings may be a Han root symbol or a sequential Han root string.A Han root string can be represented again in this way,until its Han root string is a Han root symbol or a Han root string as defined in form 2 . This can be described as:

represent

<sequential Han root string> →
 <structure symbol>
 <Han root string>
 <Han root string>

For example: Han character Sequential Han root string

晰	☐日☐木斤
圆	☐口☐口贝
擎	☐☐扌丸手
焯	☐火☐☐亻七十

2.2 The Practice

In practice, the method of describing Han characters linearly has been used since the start of the Han character creation. Since the pronunciation itself cannot uniquely identify a Han character, people often resort to the description orally(sequentially) to describe a Han character. The following provides several examples where the combining symbol method are used.

Example 1: Communication

The Technical Standard of Encoding Han Characters on radio paging which drawr up by Ministry of Posts and Telecommunications of P.R.C.

has been used in radio systems.

There are 1,500 Han characters that are used at most, 759 Han roots, 11 structure symbols, 100 bars of short sentences and other symbols in this Technical Standard. Transmitting Han characters by this Technical Standard, one character one code in the 1,500 Han characters that are used at most, and the others are represented by combining the Han roots with the structure symbols. The special advantage of this Standard is able to transmit more than 50,000 Han characters and need short time of holding communication channels (The average each Han character needs less than 13 bits).

A part of content see Appendix A.

Example 2: None Standardized Han Characters in Publication

Some Han characters described by combining Han roots in Han characters Frequency Table published and Used by People's Republic of China Xinhua Communication Office. Refer to Appendix B.

Example 3: Daily Usage by Common People

Han character '张' and '章' in Chinese surnames are the same pronunciation and are pronounced 'ZHANG'. When two strangers happen to meet together and introduce themselves each other, they will explain that their surname is Han character '张' combining Han root '弓' with '长', or the Han character '章' combining Han root '立' with '早'.

Example 4: The Computer Systems Adopting Combining Han Roots

There are many computer systems adopting combining Han roots in China:

. Mr Zhu Bangfu (Taiwan) can compose more than 50,000 Han characters adopting combining Han roots.

. DF-DOS adopting combining Han roots by Langchao Group, can deal with more than 50,000 Han characters.

2.3 Presentation of Han Characters

To present Han characters represented in a sequential Han root string, we can use any one of the following ways:

Way 1: To present sequential Han root string

Present all the graphic symbols of the structure symbols and the Han root symbols one by one.

Way 2: To present the generated combining graphic symbol

Present a generated graphic symbol in one graphic position.

3. JUSTIFICATION and POSSIBILITY

3.1 The Advantages

Advantage 1: Capability of expressing all Han characters.

. It is capable to express all Han characters beyond of the coded character set.

. It is easier to represent and display all Han characters in a small system.

Advantage 2: This code is capable to be exchanged as the self-named Han character.

Advantage 3: There are no duplicate codes in all Han characters with using this method.

Advantage 4: It is beneficial to recognize all Han characters with the optics pattern-recognition.

Advantage 5: Benefit to make the rule of collating Han characters consistent. Adding a new character by this method, this new character's sequence has been identified.

Advantage 6: Capability of adapting the development of Han characters, because new Han character can be made by using the Han root symbols.

Advantage 7: Only needing a small Han root ROM instead of using the large ROM in which more than ten thousand Han characters have been stored. Benefit to save the cost of computer system.

3.2 The Disadvantage

The length of every Han character's code is variable. (But it is permitted when using the Extension Level 2.)

The average Han root number of every Han character exceeds 2. Because of this, it can't replace Han characters in BMP.

3.3 Combining Characters have been used in UCS.

For example:

Latin characters

$$\tilde{a} \rightarrow a + \tilde{\circ}$$

0061 0303

The following combining characters have been used also in UCS:

.Indian characters

.Hebrew characters

.Arabic characters

So adding combining symbols in BMP of UCS to represent the Han characters is possibility.

4. CONCLUSION

In sum, for its series outstanding character, the method of combining Han roots must be considered in ISO/IEC DIS 10646. We propose:

(1) 2K space can be retained in BMP of UCS to contain Han root symbols and structure symbols.

(2) these symbols can be put in spare region of A Zone, for their function is just like that of English letters. If there are not enough space in A Zone, they can be put in O Zone.

Appendix A-1

Technical Standard of Encoding Han Characters
On Radio Paging
无线电寻呼汉字编码技术规范

Ministry of Posts and Telecommunications of P.R.C
中华人民共和国邮电部

1993.5

Appendix A -2

CONTENT

目录

1. Summary.....1
 概论
2. Constructure of Han character.....2
 汉字结构
3. Representation of Han character.....3
 汉字表示
4. Constructure of code,Representation of encoding
 characters and Table of encoding characters.....5
 代码结构、字符编码表示及字符代码表
5. Examples of encoding information
 and communication form.....9
 信息编码及传输格式举例

Appendix A -3

1. 主题内容与适用范围

1.1 本技术规范规定了一个寻呼系统向汉字显示寻呼机传达信息的模式。其中包括汉字信息的编码及传输格式,也规定了一般符号、数字、拉丁字母等字符的编码表示及传输格式。

1.2 本技术规范适用于我国国内公用无线电寻呼系统。

2. 引用标准

GB8521.1~8521.2 信息处理文本通信用编码字符集

3. 术语

3.1 字根

字根是从汉字中析取的最小组字元素的集合。

3.2 字根符

字根的可视表示。

3.3 字根块

字根符按一定的位置关系在平面上的组合称为字根块。

字根符是最小的字根块。

3.4 汉字结构

汉字结构表示汉字中字根块之间的相对位置关系。在本规范中,将汉字归为11种结构并用相应的结构符表示,见附录A。

3.4.1 独体结构

具有单一字根符的汉字结构称为独体结构。例如:“王”。




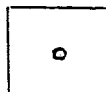



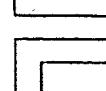
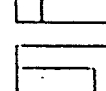
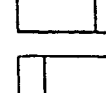
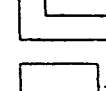
3.4.2 左右结构

两字根块在水平方向上左右排列的汉字结构称为左右结构。例如:“晒”。

3.4.3 上下结构

两字根块在垂直方向上上下排列的汉字结构称为上下结构。

Appendix A - 4

代码表上的位置 区号 位号	字 符	结构符名称	备 注
46-01		独体结构符	
46-02		左右结构符	
46-03		上下结构符	
46-04		全包围结构符	
46-05		上包围结构符	
46-06		下包围结构符	
46-07		左包围结构符	
46-08		左上包围结构符	
46-09		右上包围结构符	
46-10		左下包围结构符	
46-11		叠加结构符	

Appendix 4 - 5

29区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 一 丨 丿 ㇀ ㇁ ㇂ ㇃ ㇄ ㇅ ㇆ ㇇ ㇈ ㇉ ㇊ ㇋ ㇌ ㇍ ㇎ ㇏
 19 ㇐ ㇑ ㇒ ㇓ ㇔ ㇕ ㇖ ㇗ ㇘ ㇙ ㇚ ㇛ ㇜ ㇝ ㇞ ㇟ ㇠ ㇡ ㇢ ㇣
 20 ㇤ ㇥ ㇦ ㇧ ㇨ ㇩ ㇪ ㇫ ㇬ ㇭ ㇮ ㇯ ㇰ ㇱ ㇲ ㇳ ㇴ ㇵ ㇶ ㇷ ㇸ
 39 ㇹ ㇺ ㇻ ㇼ ㇽ ㇾ ㇿ ㈀ ㈁ ㈂ ㈃ ㈄ ㈅ ㈆ ㈇ ㈈ ㈉ ㈊ ㈋ ㈌ ㈍
 40 ㈎ ㈏ ㈐ ㈑ ㈒ ㈓ ㈔ ㈕ ㈖ ㈗ ㈘ ㈙ ㈚ ㈛ ㈜ ㈝ ㈞ ㈟ ㈠ ㈡
 59

30区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 ㇀ ㇁ ㇂ ㇃ ㇄ ㇅ ㇆ ㇇ ㇈ ㇉ ㇊ ㇋ ㇌ ㇍ ㇎ ㇏ ㇐ ㇑ ㇒ ㇓
 19 ㇔ ㇕ ㇖ ㇗ ㇘ ㇙ ㇚ ㇛ ㇜ ㇝ ㇞ ㇟ ㇠ ㇡ ㇢ ㇣ ㇤ ㇥
 20 ㇦ ㇧ ㇨ ㇩ ㇪ ㇫ ㇬ ㇭ ㇮ ㇯ ㇰ ㇱ ㇲ ㇳ ㇴ ㇵ ㇶ ㇷ ㇸ ㇹ
 39 ㇺ ㇻ ㇼ ㇽ ㇾ ㇿ ㈀ ㈁ ㈂ ㈃ ㈄ ㈅ ㈆ ㈇ ㈈ ㈉ ㈊ ㈋ ㈌ ㈍
 40 ㈎ ㈏ ㈐ ㈑ ㈒ ㈓ ㈔ ㈕ ㈖ ㈗ ㈘ ㈙ ㈚ ㈛ ㈜ ㈝ ㈞ ㈟ ㈠ ㈡
 59

31区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 ㇀ ㇁ ㇂ ㇃ ㇄ ㇅ ㇆ ㇇ ㇈ ㇉ ㇊ ㇋ ㇌ ㇍ ㇎ ㇏ ㇐ ㇑ ㇒ ㇓
 19 ㇔ ㇕ ㇖ ㇗ ㇘ ㇙ ㇚ ㇛ ㇜ ㇝ ㇞ ㇟ ㇠ ㇡ ㇢ ㇣ ㇤ ㇥
 20 ㇦ ㇧ ㇨ ㇩ ㇪ ㇫ ㇬ ㇭ ㇮ ㇯ ㇰ ㇱ ㇲ ㇳ ㇴ ㇵ ㇶ ㇷ ㇸ ㇹ
 39 ㇺ ㇻ ㇼ ㇽ ㇾ ㇿ ㈀ ㈁ ㈂ ㈃ ㈄ ㈅ ㈆ ㈇ ㈈ ㈉ ㈊ ㈋ ㈌ ㈍
 40 ㈎ ㈏ ㈐ ㈑ ㈒ ㈓ ㈔ ㈕ ㈖ ㈗ ㈘ ㈙ ㈚ ㈛ ㈜ ㈝ ㈞ ㈟ ㈠ ㈡
 59

32区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 ㇀ ㇁ ㇂ ㇃ ㇄ ㇅ ㇆ ㇇ ㇈ ㇉ ㇊ ㇋ ㇌ ㇍ ㇎ ㇏ ㇐ ㇑ ㇒ ㇓
 19 ㇔ ㇕ ㇖ ㇗ ㇘ ㇙ ㇚ ㇛ ㇜ ㇝ ㇞ ㇟ ㇠ ㇡ ㇢ ㇣ ㇤ ㇥
 20 ㇦ ㇧ ㇨ ㇩ ㇪ ㇫ ㇬ ㇭ ㇮ ㇯ ㇰ ㇱ ㇲ ㇳ ㇴ ㇵ ㇶ ㇷ ㇸ ㇹ
 39 ㇺ ㇻ ㇼ ㇽ ㇾ ㇿ ㈀ ㈁ ㈂ ㈃ ㈄ ㈅ ㈆ ㈇ ㈈ ㈉ ㈊ ㈋ ㈌ ㈍
 40 ㈎ ㈏ ㈐ ㈑ ㈒ ㈓ ㈔ ㈕ ㈖ ㈗ ㈘ ㈙ ㈚ ㈛ ㈜ ㈝ ㈞ ㈟ ㈠ ㈡
 59

Appendix A - 6

33区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 有 犬 太 歹 友 尤 巨 车 电 戈 水 戾 先 牙 瓦 止 止 支 业
 19 少 弓 日 日 日 月 中 田 円 冂 贝 内 见 内 丰 午 牛 月 手 手
 20 毛 气 壬 生 牛 升 天 女 长 尸 片 正 白 巾 尸 斤 爪 * 尸 反
 39 40 59

34区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 父 夂 冂 乏 月 丹 氏 乐 勿 风 欠 氏 衣 夕 夕 鸟 鸟 六 文
 19 方 云 干 火 为 斗 艹 义 尸 衤 穴 尤 心 丰 尸 尹 尹 尹 尺
 20 夫 矢 弔 丑 月 卍 尸 巴 址 出 世 艮 办 办 又 予 以 小 书 册
 39 40 59

35区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 水 毋 丰 夨 马 玉 挂 末 未 击 戾 示 戈 戈 正 正 甘 世 产
 19 本 木 业 市 可 丙 石 夫 龙 戊 友 平 圭 戍 东 东 卡 占 凸 步
 20 卢 卢 业 小 尚 出 且 目 申 甲 电 屯 冉 田 由 虫 史 央 目 册
 39 40 59

36区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 册 皿 冂 凹 四 回 丹 丹 丰 生 矢 失 乍 禾 丰 丘 白 斥 尸
 19 瓜 乎 彡 义 用 用 鬼 鸟 氏 勿 勾 册 鸟 包 尔 乐 主 市 广 立
 20 半 羊 头 兴 穴 衤 必 永 丰 丰 犬 以 目 民 邪 弗 正 正 出 皮
 39 40 59

Appendix A -7

37区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 发 夬 水 水 矛 母 耒 耂 戔 耑 丰 夙 戔 其 耳 聿 亚 亩
 19 再 夙 吏 而 而 束 兩 西 西 在 册 有 百 而 页 存 彥 互 成 夾
 20 夷 臣 臣 戔 至 豕 与 卢 光 且 且 卑 曳 虫 曲 曲 馬 田 肉 朱
 39 40 59

38区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 雷 缶 先 竹 竹 臼 禾 自 自 血 血 角 月 舟 兆 炎 伞 凶 义
 19 鳥 豕 彘 齐 齒 衣 亦 产 产 亥 羊 羊 羊 并 兴 米 州 良 农
 20 聿 聿 目 艮 尔 糸 丞 麦 長 走 赤 聿 其 耶 严 巫 巫 車 甫 更
 39 40 59

39区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 亞 曲 束 夷 豆 西 兩 酉 辰 百 百 戔 夾 龙 豕 来 長 鬲 臣
 19 求 戔 鹵 亞 戔 里 貝 見 串 肉 金 我 印 身 虫 兔 席 采 豕 坐
 20 谷 龟 夬 免 角 卵 言 产 辛 亥 美 羌 尚 良 兼 頁 耶 卍 亞 其
 39 40 59

40区 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19
 01 來 來 重 束 事 雨 頁 戔 豕 建 非 齿 虎 果 門 虫 屯 垂 無
 19 乖 乘 耳 央 佳 卑 阜 自 金 戔 食 鱼 兔 肉 高 卒 戔 聿 聿 聿
 20 承 亟 函 馬 甚 革 虫 虫 束 頁 面 非 是 禹 章 乘 聿 重 耳 崩
 39 40 59

Appendix A - 8

41区	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	
01 ┆ 19 ┆ 20 ┆ 39 ┆ 40 ┆ 59		鬼	禹	爰	食	奂	風	差	酋	首	為	飛	鬲	門	鬲	麥	門	乘	我	重
	鳥	高	雀	亞	幾	首	董	莫	黃	爽	鹵	雀	魚	麻	鹿	宰	帶	鼎	戠	黑
	黍	為	幾	甞	肅	爾	齊	臧	齒	鹽	禽	龜								

Appendix B

序号	库外字描述字	次数	累计	序号	库外字描述字	次数	累计
0001	容加旁	106	106	0051	入加旁	5	942
0002	容加旁加鸟旁	71	177	0052	入加月旁	5	947
0003	世加竹头	57	234	0053	票加旁	5	952
0004	冬加头	51	285	0054	步加土旁	5	957
0005	露加头	48	333	0055	贝上加	5	962
0006	蒙加旁	42	375	0056	刀加工旁	5	967
0007	布加日旁	32	407	0057	兆加页旁	5	972
0008	奇加土旁	31	438	0058	兹加口旁	4	976
0009	页加旁	30	468	0059	爵加口旁	4	980
0010	舞加旁	25	493	0060	布加口旁	4	984
0011	舞加旁右加鸟	24	517	0061	翼去加旁系	4	988
0012	溥加木旁	24	541	0062	十加去旁见加石	4	992
0013	参加木旁	21	562	0063	觉加去旁	4	996
0014	泉加木旁	19	581	0064	詹加月旁	4	1000
0015	句加旁	19	600	0065	风加旁	4	1004
0016	平加口旁	18	618	0066	奉加王旁	4	1008
0017	燕加日旁	18	636	0067	宁加旁	4	1012
0018	翼去加旁	18	654	0068	孟加旁	4	1016
0019	品下加山旁	16	670	0069	威加虫旁	4	1020
0020	厉加木旁	16	686	0070	郎加旁	4	1024
0021	罗加口旁	15	701	0071	郞去加旁	4	1028
0022	当加口旁	15	716	0072	旁加土旁	4	1032
0023	奔加石旁	15	731	0073	並加木旁	4	1036
0024	殷下加木旁	13	744	0074	爰加旁	4	1040
0025	既下加鱼旁	13	757	0075	爰去加旁加旁	4	1044
0026	既下加石旁	11	768	0076	爰去加鱼旁	4	1048
0027	广内加旁	11	779	0077	豆加旁	4	1052
0028	私加旁	11	790	0078	酉下加旁再加木旁	4	1056
0029	井加田旁	10	800	0079	木上加旁	4	1060
0030	燕加女旁	8	808	0080	青加旁	4	1064
0031	詹加山旁	8	816	0081	炎加旁	4	1068
0032	湖去加旁	8	824	0082	召加火旁	3	1071
0033	監去加旁	8	832	0083	宛加鸟旁	3	1074
0034	瀨去加旁	8	840	0084	么加口旁	3	1077
0035	小加旁	7	847	0085	堂加口旁	3	1080
0036	宗下加旁	7	854	0086	冂加口旁	3	1083
0037	巫加旁	6	860	0087	冂加旁	3	1086
0038	君上下加竹	6	866	0088	门内加旁	3	1089
0039	白下加田旁	6	872	0089	金上加山旁	3	1092
0040	匀加旁	6	878	0090	安加石旁	3	1095
0041	亨加旁	6	884	0091	朱加石旁	3	1098
0042	戴加旁	6	890	0092	盆加旁	3	1101
0043	戴加王旁	6	896	0093	高加旁	3	1104
0044	平加王旁	6	902	0094	尺加旁	3	1107
0045	耶加旁	6	908	0095	留加王旁	3	1110
0046	炎加旁	6	914	0096	朔下加王旁	3	1113
0047	堂加足旁	6	920	0097	革加旁	3	1116
0048	斌下加贝旁	6	926	0098	勾加旁	3	1119
0049	木上加旁	6	932	0099	难上加旁	3	1122
0050	努加口旁	5	937	0100	温去加旁	3	1125