CKC Chinese Input Method

纵横输入法

Esther S. C. Chan, K. H. Tse & Kevin K. F. Ng

Keyboard input methods can be classified into 3 main types: by encoding, by pronunciation, and by structure of the characters. Different people are most comfortable with different methods and each standard has its strengths and weaknesses. For example, for someone who is already familiar with Putonghua and therefore Pinyin, the Pinyin method can be learned most quickly. However, the maximum typing rate is limited, and learning the Pinyin system is not easy. Wubi takes much effort to learn, but expert typists can enter text much faster than the phonetic methods. For people who can speak the standard Putonghua, it is easy to use the Pinyin input method. But there are still many ethnic Chinese who cannot speak Putonghua, they speak with their own dialects such as Cantonese.

Character Coding

The merit of CKC Chinese Input Method is that it does not require the user to master Pinyin because it makes use of the characteristics of the square shape of the Chinese characters. It partitions each character into 4 parts: upper left corner, upper right corner, lower left corner and lower right corner. The user only needs to perceive each character as a picture and divide it into 4 parts. Then each part is represented by a digit - that means the CKC Chinese Input Method can use up to 4 digits (0-9) to represent a character.

The strokes or the shapes of the Chinese characters components are classified into 10 groups.

"1" represents a horizontal stroke.

"2" represents a vertical stroke or the similar shapes.

"3" represents a dot or a slanting downward stroke towards the right.

` ㅛ ㅗ ⌒ ㇏

"4" represents a cross or the radical of "grass".

十 乂 ナ 乂 十 艹

"5" represents a combination of strokes with two intersections.

龙 ✕ 扌 丰 于
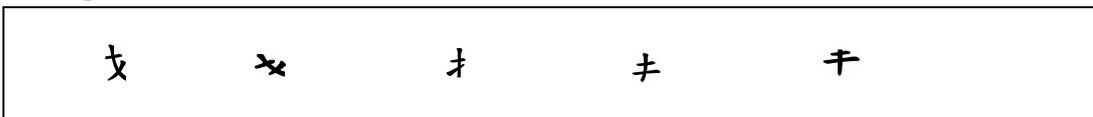
"6" represents a square.

口 口
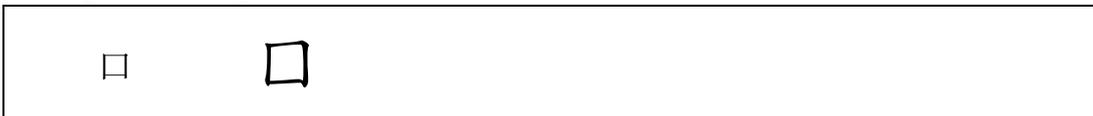
"7" represents a single corner, two upper corners, two lower corners or the shape of a hook facing right.

冂 冂 凵 ㇔ ㇙
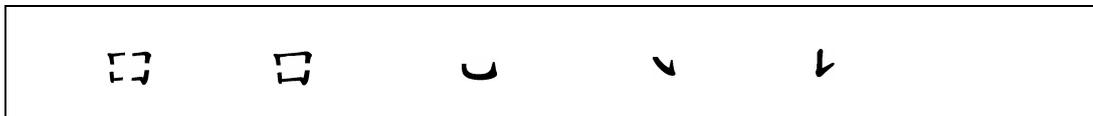
"8" refers to the shape of the number "eight" in Chinese or its shape in reverse. It also refers to the radical of "person".

八 八 丷 人 人 ㇒
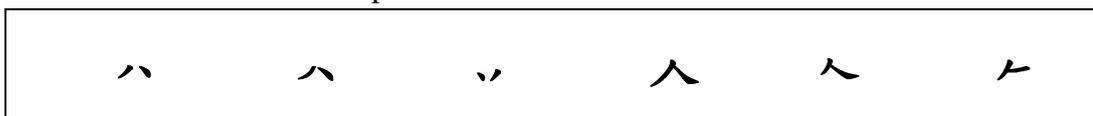
"9" represents the Chinese character which means "small" or its shape in reverse. It also refers to the radical of "upright heart".

小 小 ㅛ ㅛ 忄
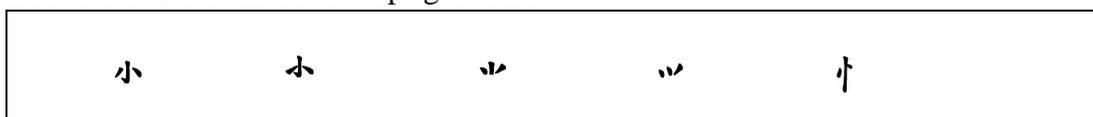
"0" represents a slanting stroke that looks like a slash or the shape of a hook facing left.

㇀ 丿 丿 亅 ㇆ ㇆
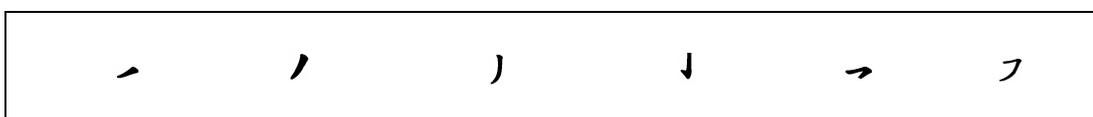
In case users forget the digit mapping, they can always click on the button in the input method bar to activate the hint. (see Figure 1)

The main principle of composing CKC code to form a single Chinese character is:

Take the stroke of the upper left hand corner as the 1st code.
Take the stroke of the upper right hand corner as the 2nd code.
Take the stroke of the lower left hand corner as the 3rd code.
Take the stroke of the lower right hand corner as the 4th code.
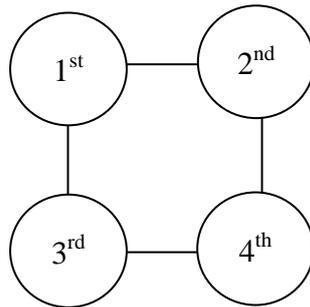
For example, the character of "陝", the 1st code is "7", the 2nd code is "4", the 3rd code is "2", the 4th code is "8". The CKC code is "7428".

7 陝 4
2   8

However, the CKC code of some Chinese characters may not necessarily be a 4 digit code. For example, the character of "西", the 1st code is "1". Since the horizontal stroke can represent both the upper left and upper right corners so it does not need the 2nd code. This rule also applies to the lower left and lower right corners, that means only 1 digit "6" is needed. For this character the CKC code is "16".

1
西
6

To input a Chinese character with the CKC Input Method, the most convenient way is to use the numeric keypad. After keying in the code, users have to complete the input with the key "slash" or "space". If more than 1 character can be represented by the same CKC code, the system will show all the characters for selection. Users can use numeric pad "0-9" for selection, and use the key "minus" for page up and "plus" for page down.

(see Figure 2) Though users are advised to use the numeric keypad for input, the IME also allows direct input using the top numeric keys on a QWERTY keyboard.

Phrase Coding

Chinese input is based on phrase as an entity. This is like typing a word in English. In English, we think of a word as an entity and we type it on the keyboard without breaking it down to alphabets. For example, for the word "today", we type the 5 alphabets naturally without thinking into the individual alphabets that forms the word. Unfortunately many popular Chinese input methods such as the Cangjie, Wubi and Q9 do not evolve to this stage yet. Though these Chinese input methods provide the association words that can form a phrase with typing the first word and then choose the other words. This kind of method is inconvenient and inefficient.

Contrary to many popular Chinese input methods, the CKC Chinese Input System allows input of phrases rather than characters to increase the input speed. To input an arbitrary phrase, we only need to specify up to 6 digits. The system has predefined more than 250,000 common Chinese phrases. The principle of coding a phrase is to iteratively apply the coding principle for each individual character that forms the phrase, from the 1st till the 5th character of the phrase. The rest of the phrase do not contribute nor affect the formation of its CKC code and therefore can be safely ignored. This implies we have limited the number of keystrokes required to input any arbitrary phrase to 6, saved lots of keystrokes and time required.

There are four variations of theme or rules that can be applied to derive the corresponding CKC code:

2-character phrase

Get the first 3 digits at most of each character. This is known as the "3 + 3" rule. The minimum number of digit of this phrase is 2 and the maximum is 6, it is because a minimum of 1 digit can represent a single character. For example the phrase "一日" (one day) requires only 2 digits , that is "16". After entering the CKC code, users have to complete the input with the key "asterisk". If a phrase has the maximum of 6 digits, the completion key can be saved.

3-character phrase

Get the first 2 digits at most of each character. This is known as the "2 + 2 + 2" rule. The minimum number of digit of this phrase is 3 and the maximum is 6. For example the phrase "一定会" (surely will) only 5 digits are needed, that is "13887".

4-character phrase

Get the first 2 digits at most of the 1st character, then the first digit of the 2nd and the 3rd characters. Finally get the first 2 digits at most of the 4th character. This is known as the "2 + 1 + 1 + 2" rule. The minimum number of digit of this phrase is 4 and the maximum is 6. For example the phrase "一试就知" (you will know immediately with a trial) only 5 digits are needed, that is "13386".

Long phrase

Any phrase which is composed of more than 4 characters is considered as a long phrase. Get the first 2 digits at most of the 1st character, then the 1st digit of the 2nd, 3rd, 4th and 5th characters, the rest of the phrase are ignored. This is known as the "2 + 1 + 1 + 1 + 1" rule. The minimum number of digit of this phrase is 5 and the maximum is 6. For example the phrase "中华人民共和国" (the People's Republic of China) can be entered with just 5 digits "50874".

From the above illustration, you can see that if users input "中华人民共和国" one by one character, they have to type in 5, 044, 8, 777, 48, 069, 6 with number of strokes for slash key and selection in between. Now, we just need to key in "50874" and the asterisk key, the phrase can be entered successfully.

Other Features

Besides providing a large number of predefined phrases, the CKC Chinese Input System also allows users to define their own phrases, such as "陕西师范大学国际汉学院" (International College of Chinese Studies, Shaanxi Normal University) with any code they like. Users just need to click on the button in the input method bar to activate the "Phrase builder", then they can define any phrase up to 40 Chinese characters with any code they choose. (see figure 1)

On top of that, there are many other features in CKC Chinese Input System to help the users to input Chinese characters more easily and quickly, such as the quick key (快速键), wild card key (万能键) and associated key (联想键). (see figure 2)

For more detail of the input method, please visit the homepage of "纵横码" at http://www.ckcsys.com.hk, "CKC Center for the Development of Information Technology in Chinese Language Teaching" at www.ckcitlang.net and try the "CKC Chinese Dictionary 纵横码字典 V0.02" at ckc.ied.edu.hk/

Figure 1

Input mode select: "A" to input digits, otherwise input Chinese — 英文状态 / 中文状态

Frequency modulation — 字词的频度调整功能

Phrase builder — 自定义功能

Half moon (numbers and alphabets in English style)
Full moon (numbers and alphabets in Chinese style) — 半角 / 全角 — 数字/英文字形选择

软键盘 — Open / close soft keyboard — 打开 / 关闭

显示纵横码的十种笔形 — CKC digit mapping

全角中文标点 — Punctuation style select — 打开 / 关闭

Figure 2

Character completion & 2-character phrase associated key

Phrase completion & 3/4-character phrase or long phrase associated key

Wild card key — 非常用单字输入切换 万能键

单字结束键 / 二字词联想键

词组结束键 三、四、多字词联想键

标点符号输入键 向前翻页键 — Page up / Punctuation

输入码 码元

向后翻页键 记忆键 — Page down

输入码全部清除 换行键 — Delete all input codes

删除光标所在位置前的一个中文字

输入码删除键 — Delete the last input code

Reference:

Chinese Input Methods for Computers -

http://en.wikipedia.org/wiki/Chinese_input_methods_for_computers

CKC Chinese Input System - http://en.wikipedia.org/wiki/CKC_Chinese_Input_System