

## **A Chinese Character Input Method**

### **Technical Field**

The invention relates to a Chinese character encoding method and a Chinese character input method, more specifically, relates to a Chinese character input method used in the computer system and its related Chinese character encoding method.

### **Background of the Invention**

Chinese computers have been popular for more than a quarter of a century, and various Chinese character input methods have been circulated in the market, they are numerous and diversified. For example, some Chinese character input methods are based on handwriting, their recognition rate is generally low, and cannot be used as fast input. Some input methods are based on Chinese Pinyin, but due to large number of homophones in the language, the steps of selecting characters will slow down the input speed.

### **Summary of the Invention**

Thus, the example of the invention provides a new Chinese character input method to overcome above technical problems.

In one respect, the invention provides a Chinese character encoding method which comprising following steps: Decompose the Chinese characters or words needing to be input into one or more font section(s) based on the decomposition rules, wherein each of said font sections comprises: at least one basic character; and based on the code selection rules, respectively select a corresponding character code to represent each of said font sections according to said basic characters contained therein; and based on the encoding rules, combine the selected corresponding character codes into character code combinations, to represent said Chinese characters or words needing to be input. Wherein, said font section comprises: based on said

decomposition rules, said Chinese characters needing to be input or several Chinese characters of said words needing to be input that do not need to decompose or cannot decompose; wherein said word comprises: respectively combined by two or more Chinese characters.

Preferably, said basic character comprises: it is specified as the smallest unit in the font, it cannot be decomposed, and it is set to combine with other basic characters to become all Chinese characters.

Preferably, said basic character comprises: classify into nine types according to font structures, comprising 4-side character, 3-side character, 2-side character and 1-side character that defined by the number of sides, 1-stroke character, 2-stroke character, 3-stroke character and 4 stroke character that defined by the number of strokes, and intersecting character.

Preferably, said 4-side character comprises: the character roughly displayed as a square "□";

Preferably, said 3-side character comprises: the character that any one side of the square "□" is missing.

Preferably, said 2-side character comprises: the character that any one side of the 3-side character is missing, comprising the character roughly having one angle of square or rhombus, the character having an acute angle or an hook shape, wherein said square angle, rhombus angle, acute angle and hook shape represent as additional descriptions of said 2-side character.

Preferably, said 1-stroke character, 2-stroke character, 3-stroke character and 4-stroke character comprise: the horizontal, vertical, slant and/or dot stroke, having 1-4 stroke(s) that do not intersect with each other, link together, and do not intersect with other strokes at all or all intersect with one or more other strokes; wherein said horizontal, vertical, slant and dot represent as additional descriptions

of said 1-4 stroke characters; wherein a horizontal, a vertical and a slant stroke of said 1-stroke character are equivalent to said 1-side character.

Preferably, said intersecting character comprises: the character which only comprising two intersecting lines.

Preferably, said basic character further comprises: at least a specially defined character.

Preferably, said 4-side character comprises: when one of the strokes of the character "□" extending to outside "□", said extended stroke is separated on the edge of said character "□", and respectively belongs to said character "□" and other basic characters; wherein said strokes belonging to outside said character "□" represent the said additional description of said 4-side character.

Preferably, said 2-side character comprises: when a single stroke has more than one turn (angle), said multiangular stroke is separated on said turn, and respectively belongs to a 2-side character and other basic characters.

Preferably, said basic character comprises: the single stroke cannot belong to two or more basic characters except said extended stroke and multiangular stroke of character "□".

Preferably, said basic character comprises: a said basic character or a said basic character with additional description is represented by a letter feature; wherein, the same basic character with different additional descriptions will be represented by different letter features; wherein, said additional description comprises the stroke structure rules of said basic character or/and the stroke structure rules of outside said basic character.

Preferably, said letter feature comprises: classified into ten groups, the basic character that represented by each said letter feature in the same group having the same number of sides or/and strokes.

Preferably, said letter features in each group comprise: contain no more than two types of said basic characters, different said letter features in the same group is composed of said no more than two types of basic characters and different additional descriptions.

Preferably, said letter features in each group comprise: one or more said letter feature(s) in the same group is(are) represented by a said corresponding character code.

Preferably, said corresponding character code comprises: each basic character in each said font section, based on said letter feature represented, is corresponding to one said corresponding character code.

Preferably, said letter features in each group comprise: code selection priority level of said letter features in each group from low to high is represented by the number 1 to 10.

Preferably, said ten priority levels of code selection comprise: the more number of sides or/and strokes of said basic character that represented by said letter features in each group, the higher the code selection priority level will be.

Preferably, said ten priority levels of code selection comprise: each said corresponding character code corresponds to one priority level of code selection according to said letter feature represented.

Preferably, said code selection rules comprise: compare the code selection priority level of letter feature to which each basic character in said font section belongs, wherein the corresponding character code to which said letter feature belongs with highest level is selected to represent said font section.

Preferably, said code selection rules comprise: when there are more than one said letter feature in highest priority level of code selection, wherein select the one in

highest level, and the corresponding character code of said letter feature of basic character represented with first writing stroke order to represent said font section.

Preferably, said decomposition rules comprise: according to the complexity of decomposition steps, said decomposition rules are classified from simple to complicated as: 2-code mode, general mode, half-element mode and over half-element mode; said modes of decomposition rules further comprise word mode.

Preferably, said decomposition rules comprise: based on the specified steps of one of said modes of decomposition rules, and based on the decomposition method of said decomposition rules to decompose said Chinese character or word needing to be input into one or more said font section.

Preferably, said each mode of decomposition rules comprises: when each of said font sections comprises more than one basic character, and needs to be decomposed into two relatively smaller font sections according to the regulations of said decomposition rules mode, then decompose said font section into two relatively smaller font sections based on decomposition method; and repeat until finishing the specified steps of said decomposition rules mode; when said font section specified to be decomposed has only one basic character that cannot be decomposed any more, then ignore said step;

Preferably, the decomposition steps specified in said 2-code mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method.

Preferably, the decomposition steps specified in said general mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method, then further decompose each of acquired font sections into two smaller font sections respectively based on said decomposition method.

Preferably, the decomposition steps specified in said half-element mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method; when one of two acquired font sections is half-element character and said half-element character locates in specified position in the Chinese character needing to be input, then further decompose another font section outside said half-element character into no more than three font sections based on specified decomposition steps of the section outside half-element character.

Preferably, the decomposition steps specified in said over half-element mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method; when one of two acquired font sections is over half-element character and said over half-element character locates in specified position in the Chinese character needing to be input, then further decompose said over half-element character into two smaller font sections based on said decomposition method; and further decompose another font section outside said over half-element character into no more than three smaller font sections based on specified decomposition steps of the section outside said half-element character.

Preferably, said half-element character comprises: the thirty-four most frequently occurring radicals or component characters in Chinese characters, and the characters which are similar to them;

Preferably, said over-half-element character comprises: ten characters in said half-element characters.

Preferably, said half-element mode or over half-element mode comprises: said specified position of half-element character or over half-element character is where said half-element character most frequently occurring in Chinese characters.

Preferably, wherein the decomposition steps specified in said section outside half-element character comprise: decompose said another font section outside half-element character or over half-element character into two smaller font sections

based on said decomposition method; in two acquired smaller font sections, when the one with first writing stroke order is also a half-element character or only comprises a basic character, then further decompose the one with latter writing stroke order into two smaller font sections based on said decomposition method; and when said one with first writing stroke order is not a half-element character or said one with latter writing stroke order only comprises a basic character, then further decompose said one with first writing stroke order into two smaller font sections based on said decomposition method; wherein, the decomposition steps specified in said section outside half-element character further comprise: further decompose said another font section outside half-element character or over half-element character into two smaller font sections based on said decomposition method; and further decompose each of acquired font sections into two smaller font sections respectively based on said decomposition method; and when more than three font sections acquired from above steps, then compare the code selection priority level corresponding to corresponding character code for each of the font sections acquired from above steps, remove one said font section which has lowest priority level of code selection and the writing stroke order of its corresponding basic character is latter, will not be used as said font section for encoding.

Preferably, the decomposition steps specified in said word mode comprise: when it is a 2-character word needing to be input, decompose the first and second character of said 2-character word into two font sections respectively based on said decomposition method; when it is a 3-character word needing to be input, decompose the first character of said 3-character word into two font sections based on said decomposition method, and use the second and third character of said 3-character word to act as a font section respectively without decomposition; when the word needing to be input comprises more than three characters, use the first, second, third and last character of said word to act as a font section respectively without decomposition, the other characters will not be used as said font section for encoding.

Preferably, said decomposition method comprises: preferably decompose the said Chinese character needing to be input or the font section needing to be further

decomposed into two font sections which respectively comprising a character, radical or common character.

Preferably, said decomposition method comprises: preferably decompose the said Chinese character needing to be input or the font section needing to be further decomposed into two font sections that the size is similar or the shape is square.

Preferably, said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed comprises the sections separated from each other, preferably decompose said Chinese character or said font section into two font sections that were previously separated from each other.

Preferably, said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed does not comprise the sections separated from each other, preferably decompose said Chinese character or said font section into two font sections that no stroke intersecting with each other in said Chinese character or said font section.

Preferably, said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed only has the section with intersecting strokes, preferably decompose said Chinese character or said font section into two font sections that one stroke is horizontal, another one stroke is vertical.

Preferably, said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed has similar section with other Chinese character or other font section, preferably decompose said Chinese character or said font section into two font sections according to their differences.

Preferably, said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed has the

structure that one character clamping another character, preferably decompose said Chinese character or said font section into two font sections which are dominant character and clamped character.

Preferably, said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed does not comprise the sections separated from each other, preferably select to separate an extended stroke or multiangular stroke of said character "口", decompose said Chinese character or said font section into two font sections.

Preferably, said encoding rules comprise: each of said corresponding character codes in said character code combination is arranged according to writing stroke order of its corresponding basic character.

Preferably, said corresponding code comprises: comprising different letter modes, each said letter mode uses different numbers of corresponding character codes to represent different letter features in each said group of letter feature.

Preferably, said letter mode comprises: ten letter modes comprise ten corresponding character codes, respectively use one said corresponding character code to represent different letter features in each group.

Preferably, said letter mode comprises: forty letter modes comprise forty corresponding character codes, respectively use five, four or three said corresponding character codes to represent different letter features in each group.

Preferably, said letter mode comprises: twenty six letter modes comprise twenty six corresponding character codes, respectively use three, two or one said corresponding character code to represent different letter features in each group.

On the other hand, the invention provides a Chinese character input method comprising: receive the input code combination that input by keyboard; match the input code combination with said character code combination that representing each Chinese character or word in the Chinese character encoding table established

according to said Chinese character encoding method; and display the Chinese characters or words that represented by said input code combination.

Preferably, said input character code comprises: the character code combination that completely matching with the record of said Chinese character encoding table.

Preferably, said input character code comprises: the character code combination that partially matching with the record of said Chinese character encoding table.

Preferably, said displaying the Chinese character that matching with said input code combination comprises: display all Chinese characters or words which at least partially matching with said record in a text index.

Preferably, said displaying the Chinese character that matching with said input code combination comprises: when need to display several Chinese characters in said record that matching with said input code combination, the arrangement order of said each Chinese character is displayed from low to high according to code selection priority level for the letter features represented by the corresponding character codes in character code combination encoded in said each Chinese character.

Preferably, said receiving the input code combination that input by keyboard comprises: each key that representing said input character code on the keyboard corresponds to a said corresponding character code with the same name.

Preferably, said key comprises: the position of said corresponding character code that corresponding to it on said keyboard is arranged in order from 1 to 10 according to code selection priority level of letter feature represented.

Preferably, said key comprises: the position of said corresponding character code that corresponding to it on said keyboard is arranged in the order horizontal, vertical, slant and dot according to additional description contained in letter feature represented, wherein the ones with same type of additional descriptions are in adjacent positions.

Preferably, said keys representing each of input character codes comprise: in specific mode, one said key corresponds to two said corresponding character codes; selecting/pressing said key is equivalent to inputting said two corresponding character codes.

Preferably, said keyboard comprises: the tools used for inputting messages in electronic system/device.

Preferably, said keyboard comprises: the manner that on said keyboard can correspond to said corresponding character code.

On the other hand, the invention provides a Chinese character encoding method which comprising following steps: decompose the Chinese characters needing to be input into one or more font section(s) based on the decomposition rules, wherein each of said font sections comprises at least one basic character; based on the code selection rules, select the character code corresponding to one basic character in said each of font sections; and based on the encoding rules, combine the selected character codes into character code combinations, to represent said Chinese characters needing to be input; wherein, when each of said font sections can be further decomposed based on said decomposition rules, then decompose each of said font sections into two smaller font sections respectively based on decomposition rules.

Preferably, said at least one basic character comprises ten groups of letter features, each group of letter features is represented by one or more said corresponding character code(s).

Preferably, wherein said at least one basic character further comprises specially defined character.

Preferably, said corresponding character code comprises: 0-9 digit keys on the computer keyboard representing said each group of letter features.

Preferably, said corresponding character code further comprises the characters below 0-9 digit keys on the computer keyboard.

Preferably, said each group of letter features is represented by one digit of 0-9 keys on the computer keyboard and one to three character(s) right below said digit on the computer keyboard, wherein said digit and said one to three character(s) right below respectively represent different characters in the same group of letter features.

Preferably, each of said multiple half-element characters represents the common character structure of the Chinese character structure.

Preferably, said encoding rules comprise: use a character code or 2-character code combination to represent said half-element character.

Preferably, said encoding rules comprise: use 3-character code combination to represent remaining section outside said half-element character in the Chinese character.

Preferably, said decomposition rules comprise decomposing said Chinese character needing to be input into two font sections, and when each of font sections can be further decomposed based on said decomposition rules, then decompose each of font sections into two smaller sections based on said decomposition rules.

Preferably, each said smaller section is represented by a single character code.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed is composed of two characters, radicals and/or common characters, then decompose said Chinese character or said font section into two sections.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed is composed of

two similar size sections, then decompose said Chinese character or said font section into two sections.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed is composed of two sections separated from each other, then decompose said Chinese character or said font section into two sections.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed has similar section with other Chinese character or other font section, then decompose said Chinese character or said font section into two sections according to their differences.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed has the font section that one is clamping another one, then decompose said Chinese character or said font section into two sections.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed has a single stroke that containing more than one turn, then decompose said Chinese character or said font section into two sections.

Preferably, said decomposition rules comprise when said Chinese character needing to be input or the font section needing to be further decomposed has obvious said at least one basic character and linked other font section, then decompose said Chinese character or said font section into two sections.

Preferably, said decomposition rules further comprise: using a single character code to represent the Chinese character or font section that cannot be decomposed.

Preferably, said based on encoding rules further comprise selecting said character code combination according to code selection priority of said letter feature.

Preferably, said code selection priority comprises: prefer to select the character code that representing relatively complicated basic character.

Preferably, said code selection priority comprises: select the character code according to writing stroke order of said font section.

Preferably, said encoding rules comprise: said character code combination is limited to contain one or two character code(s).

The input method of the invention has a lot of advantages. Wherein, the Chinese character encoding method used by the input method "extracts" some of the simplest and the most easily recognizable structures -- use the number of sides (4-side to 1-side), number of strokes (1-stroke to 4-stroke) and intersecting character to define basic character. Each letter basic character will comprise only one of above three feature structures, will not mix with other strokes or characters, thus it can be clear at a glance.

As said basic character comprises the features which can be arranged in a sequence, and other rules of the input method are also set based on this property, it is of great help to the learning and use of the input method.

Besides, the Chinese character encoding method of the invention uses the method of selecting the complexity and discarding the simplicity (the code is selected from the highest level of the letter feature to which the character belongs, and the level will be defined with reference to the number of sides and strokes of the character). The advantage is that the larger the number of sides or strokes, the more obvious the structure will be, the more prominent it becomes in the font, which becomes to encode by selecting the most easily recognizable section of the font. At the same time, also in accordance with other rules of the invention, it is often possible to select the difference between different words for encoding, reducing the chance of repeated code.

#### **Brief Description of the Drawings**

The performance and advantages of the invention can be further understood with reference to the rest content and the drawings; the same component/section in these drawings has the same code. In some cases, sub-code is placed after a code and a hyphen to indicate one of many similar components/sections. When a code is mentioned but without specific sub-code, it refers to all these similar components/sections.

Figure 1 shows an example of a Chinese character encoding method of the invention. Figure 2 shows different decomposition possibilities of character "口" and examples of code selection priority, and keyboard layout, letter features and basic characters of "Chinese Code";

Figure 3 shows letter features of "Goldstone Code";

Figure 4 shows the basic characters and letter features of Chinese character encoding shown in Figure 1, and ten groups representing the priority of code selection;

Figure 5 shows the keyboard layout of "Goldstone Code";

Figure 6 shows the half-element characters of "Goldstone Code" and "half-element method";

Figure 7 shows the keyboard layout of "Goldstone" shown in Figure 5 developing into the keyboard layout of "26-Code"; and

Figure 8 and Figure 9 show the keyboard layout and simplified style half-element characters when using "Goldstone Code" with "simplified Chinese character".

### **Detailed Examples of the Invention**

The different advantages and strengths provided by various examples of the invention can be easily understood from the following descriptions.

The inventor believes that a successful input method needs to meet the following two conditions: (1) Easy to learn and master. (2) Efficient. It is not very difficult to meet one of the above requirements, but as the Chinese font structure is extremely

complex and changeable, and the font style is rich and diverse, it is not an easy task to balance the both requirements and find the "optimized" form therein.

Generally speaking, the Chinese character input method can be divided into two categories. One has a small number of letters, which is relatively easy to learn, but it is difficult to avoid selecting words and cannot achieve maximum efficiency. Another one has more letters and higher efficiency, but it is generally more difficult to master. The former is suitable for small-sized device, such as mobile phones; the latter can be used with a keyboard on a computer to satisfy the users who have higher requirements for input speed. Above two types of input methods have their own advantages, but it is a bit difficult for users to master two different input methods at the same time.

Referring to Figure 1, the invention provides a Chinese character encoding method, comprising the following steps: Decompose the Chinese characters or words needing to be input into one or more font section(s) based on the decomposition rules, wherein each of said font sections comprises: at least one basic character; and based on the code selection rules, respectively select a corresponding character code to represent each of said font sections according to said basic characters contained therein; and based on the encoding rules, combine the selected corresponding character codes into character code combinations, to represent said Chinese characters or words needing to be input. Wherein, said font section comprises: based on said decomposition rules, said Chinese characters needing to be input or several Chinese characters of said words needing to be input that do not need to decompose or cannot decompose; wherein, said word comprises: respectively combined by two or more Chinese characters.

As shown in Figure 1, the Chinese character encoding method of the invention can "translate" ideographic characters, especially Chinese characters, into character code combinations, for example, the Chinese character "解" (102) can be converted into a code combination (4762) (104) through the Chinese character encoding method (100). Since each code in (4762) can be directly input to the calculator/computer device by a computer keyboard, establishing a Chinese

character encoding table with different ideographic characters and corresponding character codes, so that used as the input method of computer/calculator. When using the input method, the user will first convert the Chinese characters needing to be input into the code that can be directly input by the keyboard based on certain rules, and then call out the Chinese characters needing to be input by typing the correct code on the device and "Enter" into the device.

Above example of code selection rules described how to arrange letters into different codes according to different structures of each character. After the letters are assigned to the keys on the keyboard, the corresponding characters can be input by typing codes.

The correspondence between the code and the character is the basis of the input method. As long as each character is paired with a good code, the third-party software such as Microsoft's "Universal Input Method Editing Tool" and "Lime Input Method" can be used to create a half-split input method and use in the computer or mobile phone. Or, a set of proprietary program can better play the effectiveness of the input method.

For example, user may execute a Chinese character input method in a computer device to input the Chinese characters needing to be input. The computer device receives the input code combination input by the keyboard, and then matches the input code combination with the Chinese character encoding table established by the Chinese character encoding method of the invention; and displays the Chinese characters matching with said input code combination. The input method of the invention may be named "half-split input method" and is further explained in the following examples.

Preferably, the decomposition steps specified in said general mode of the decomposition rules comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method, then further decompose each of acquired font sections into two smaller font sections respectively based on said decomposition method. For example, as shown in the character "解"

(102) in Figure 1, the font of a character can be decomposed into two parts like cutting fruit to get the left (102A) and right (102B) two font sections, and then decompose two acquired sections into upper and lower sections respectively to obtain four sections (may be called smaller sections) of 102A-I, 102A-II, 102B-I and 102B-II. Respectively pick a code based on the stroke structure of these four sections and the code selection rules, and according to the writing stroke order, the character code combination can be got based on the encoding rules.

In one example, "basic character" is defined as the smallest units that making up various fonts (comprising nine simplest character structures). It can no longer be decomposed, only one letter can be taken (equivalent to the corresponding code). Wherein, a letter represents one or more "letter feature(s)", each "letter feature" consists of a "basic character" or a "basic character" with additional description. Therefore, each "basic character" can correspond to a letter representing it through its "letter feature".

Preferably, said decomposition rules comprise: when each of said font sections comprises more than one basic character, and needs to be decomposed into two relatively smaller font sections according to the regulations of said mode of decomposition rules, then decompose said font section into two relatively smaller font sections based on said decomposition method. This step can be repeated again and again until finishing the specified steps of said mode of decomposition rules. Some Chinese characters may have only one "basic character" after decomposition (or the whole character), which cannot be decomposed any more, so the codes of those characters will have less than four letters. Said decomposition rules further comprise: when said font section specified to be decomposed or said Chinese character needing to be input has only one basic character that cannot be decomposed any more, then skip said step.

As far as the writing stroke order is concerned, if multiple strokes of a "basic character" are mixed with the strokes of other sections, their first strokes should be taken as the order of writing stroke order. For example, the character "田", in which

the last stroke of "口" is after the character "十", but the order should be "口" and "十" when taking the code.

In general, a Chinese character can be easily decomposed into two sections by intuition alone. As for further decomposing into four, this is just decomposing into two sections again. For some people, clear guidelines may be preferred. The inventor summed up six rules as the method to decompose the font. They are character, half, split, same, pair and cut. Wherein, "character, half, and split" are mainly defined according to intuition, which is the majority; while "same, pair, and cut" are the rules determined according to needs, and are a few special cases.

Preferably, said decomposing method comprises: Preference is given to decompose the said Chinese character needing to be input or the font section needing to be further decomposed into two font sections that each containing a Chinese character, a radical or a common character. "Character" When a font is composed of two different (can be the same) characters, it is the easiest to split them. For example, "明" can be decomposed into "日 and 月". This is also the most common case. However, when using the "character" in "character" to decompose the font into two sections, it should be noted that some radicals and common characters, such as "亠, ㇀, ㇁", etc., are also included in the definition of "character".

Said decomposition method further comprises: Preference is given to decompose the said Chinese character needing to be input or the font section needing to be further decomposed into two font sections with similar size or square shape. "Half" Select the way that can decompose the font into two sections which are relatively equal, symmetric and regular, comprising: 1. the size of two sections is similar; 2. can be decomposed by simpler dividing line (straight line). For example, "高" should be decomposed into "亠 and 同", but not "亠 and 高" or "高 and 口".

Said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed comprises the sections separated from each other, preferably decompose said Chinese character or said font section into two font sections that were previously separated from each other;

and when the sections are not separated from each other, preferably decompose said Chinese character or said font section into two font sections that no stroke intersecting with each other in said Chinese character or said font section; and when only has the section with intersecting strokes, preferably decompose said Chinese character or said font section into two font sections that one stroke is horizontal, another one stroke is vertical. "Split" Naturally, the two completely separated sections of the font will be selected to split. Then select the sections that are only in contact with each other, but do not overlap each other. The stroke structure that overlap together are usually the last choice. For example, "隶" should be decomposed into "𠂇" and "𠂆", but not "𠂇" and "𠂆".

Said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed has similar section with other Chinese character or other font section, preferably decompose said Chinese character or said font section into two font sections according to their differences. "Same" Please observe the groups "虎, 虚, 虑" and "塞, 寨, 赛" which have very similar characters, in order to avoid their encoding is the same, it is preferable to split them according to the differences. For example, "虎" should be decomposed into "虍, 几", "虚" should be decomposed into "虍, 业", and so on.

Said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed has the structure that one character clamping another character, preferably decompose said Chinese character or said font section into two font sections which are dominant character and clamped character. "Pair" The result of this decomposition method will make one section sandwich another section. For example, "街" should be decomposed into "行, 圭", "瓣" should be decomposed into "辵, 瓜" and so on. Please note that, the encoding order must still be arranged in stroke order. For example, the encoding order of "街" should be "彳, 土, 土, 予".

The fonts belong to "same" and "pair" decomposition method can usually be interpreted by the definition of "character". However, their dividing lines are very

irregular and difficult to identify, so the other types are set up to illustrate. Besides, they are not many which about twenty groups each.

Said decomposition method comprises: when said Chinese character needing to be input or the font section needing to be further decomposed does not comprise the sections separated from each other, preferably select to separate an extended stroke or multiangular stroke of said character "口", decompose said Chinese character or said font section into two font sections. "Cut" The Chinese characters are basically composed of different strokes, this is deeply ingrained in user's mind. Therefore, should try to avoid forcibly cutting a single stroke. However, in half-split input method, it still needs to cut a stroke to get reasonable code in the following two cases.

Said 2-side character comprises: when a single stroke has more than one turn (angle), said multiangular stroke is separated on said turn, and respectively belongs to a 2-side character and other basic characters.

There are very few fonts with more than one turn (corner position) in a single stroke. Such as "乚" in the character "乃". It needs to be cut at the "corner" to get two characters "乚", "丿" which meet the definition of the "letter feature" of the

Said 4-side character comprises: when one of the strokes of the character "口" "extending to outside "口", said extended stroke can be separated on the edge of said character "口", and respectively belongs to said character "口" and other basic characters.

The "口" shape font is very common in Chinese characters. It is also a very important element in the "letter feature" of half-split input method. While some font structures have obvious "口" feature, they have a long tail (and only one). For example, "巳, 尸, 艮", etc. In these cases, "cut" method needs to use to split them into characters such as "口 and 乚", "口 and 丿", "目 and 丩", etc.

The rule of "only one cutting" must be kept in mind, because cutting twice to obtain a "口" shape is absolutely not allowed. For example, the Chinese character "巨" must be decomposed into "匚, 丩", and "且" must be decomposed into "冂, 一".

The purpose of the decomposition rules is to allow users decompose the font in the simplest and intuitive way. Although the three rules of "character, half, and split" can often obtain the same decomposition result, they sometimes conflict with each other. At this time, based on above principles and select the most easily distinguishable way to split them; the input method will be optimized for the most easily distinguishable way, but the rest methods also can usually be used.

Below is the "letter features" of each letter defined by the half-split input method (ten letter modes).

Preferably, each said letter feature comprises: classified into ten groups, each of said groups comprises no more than two types of basic characters, different said letter features in the same group can be composed of said no more than two types of basic characters and different additional descriptions, and the basic character that represented by each said letter feature in the same group having the same number of sides or/and strokes. Said letter mode comprises: ten letter modes comprise ten corresponding character codes, respectively use one said corresponding character code to represent different letter features in each group. The half-split input method (10-letter mode) has ten letters. By default, the numeric keys on the top of keyboard are used as keys. According to the position on the keyboard, from right to left, the letter names and key names (equivalent to the corresponding character codes) are: "口 = (0)", "中 = (9)", "日 = (8)", "月 = (7)", "心 = (6)", "人 = (5)", "手 = (4)", "十 = (3)", "華 = (2)", "一 = (1)". In order to explain conveniently, they will be referred by their key names instead of their Chinese letter names in the following. For example (0) means the letter "口", (9) means the letter "中", and so on. Said Chinese character input method also includes: a key on the keyboard representing each input code

corresponds to a corresponding character code with the same name.

Said 4-side character comprises: the character roughly displayed as a square character "口". "口" shape is the basis of letter feature of the half-split input method. The letter features of three letters (0, 9, 8) are "口" shape with four sides. Their difference is whether there are other strokes in "口". There is not anything in "口" of (0), such as "口, 巳, 尸", etc. There will be one or two horizontal in "口" of (8), such as the characters "目, 目". Those with other character structures in "口" are classified as (9), such as "中, 田, 四, 巴", etc. Please note these three letters are only the frame of "口", the rest outside the frame can be split out and used for encoding. For example, the codes for the character "回" is (90), but not just (9). Wherein, the definition of the letter feature refers to the section other than the character "口", it is called an external feature and belongs to a form of said additional descriptions of basic character.

Said 3-side character comprises: the character that any one side of the character "口" is missing. Remove a side of "口" to get "凵", it has three sides. This is the letter feature of (7). There are four directions "凵, 匚, 凵, 凵", the strokes on both (left and right) sides can be slight slant or a hook. For example, the Chinese characters "安, 用, 凡, 且, 卯" (upper left section), "尹, 山, 甘", etc all comprise the letter features that belonging to (7).

Said 2-side character comprises: the character that any one side of the 3-side character is missing, comprising the character roughly having one angle of square. For example, remove one side of the character "凵" to get "乚" with two sides. This is the letter feature of (6). There are also four directions, and can be a slight slant or a hook, such as "厂, 乚, 丿, 丿, 乚", etc. Besides, for some characters like "乚" in Chinese character "心", "乚" in "方", their angles may not be so obvious in some writing styles, but they are still the member of letter (6).

Said 2-side character further comprises: the character roughly having one angle of rhombus, the character having an acute angle or an hook shape. Wherein said

square angle, rhombus angle, acute angle and hook shape represent as additional descriptions of said 2-side character. 2-side character can also be called angle shape. The letter features of (5) and (4) are angle shape. In comparison, the four forms of letter features of (6) like different angles on a square "┌┐" in different directions. The letter features of (5) like different angles on a rhombus "<^>" in different directions. For example, "人, 人, 人, 人" (right section), etc. However, based on the real situation of Chinese character, (5) only has the characters with upward and leftward two directions (in terms of the sharp corner). In the half-split input method, such as "丿, 丶", etc will be regarded as two dots, it will be illustrated later. The angle shape feature of (4) is acute angle or hook shape. There are four forms, comprising "ㄥ, ㄣ, ㄤ, ㄥ". Wherein the stroke with a hook can be slight slant. For example, the related sections of the characters "彡, 彡, 彡", etc.

Said 1-stroke character, 2-stroke character, 3-stroke character and 4-stroke character comprise: the horizontal, vertical, slant and/or dot stroke, having 1-4 stroke(s) that do not intersect with each other, link together, and do not intersect with other strokes at all or all intersect with one or more other strokes. Wherein said horizontal, vertical, slant and dot represent as additional descriptions of said 1-4 stroke characters. Wherein a horizontal, a vertical and a slant stroke of said 1-stroke character are equivalent to said 1-side character. The letter features of one side falls on the letter (1). A dot is included in it, so a horizontal "一", a vertical "丨", a slant "丿" and a dot "丶" all belong to the letter (1). The letter features of (2) are two horizontals, two verticals, two slants and two dots. The characters comprise: "二, 介" (lower section), 八, 勿 (middle section), 木 (two dots on left and right), 丿, 丶, etc.

Said intersecting character comprises: the character which only comprising two intersecting lines. The intersecting characters are arranged in the letter (3). The letter features are two intersecting lines which comprise three forms: "十, 乂, 卄". The structure of two intersecting lines must be definitely clear without intersecting with other strokes. For example, the Chinese character "卄" can only be split into

(12), "市" should be decomposed into (171).

Above can be referred to as the main feature of each letter. As there is an extended section in letter feature (equivalent to 1-4 stroke character). This means the mode of letter feature of (1) and (2) can be extended to other letters. This makes the letter (3) also comprise the letter features of three horizontals, three verticals, three slants and three dots. For example, "三, 川, 彡, 彳", etc. The letter (4) comprises four verticals and four dots. The related sections of the Chinese characters "杰, 雨, 氷" meet the definition of four dots. The only one example of four verticals is "舞". According to the actual situation of Chinese character, there is no five horizontals, six dots and etc character. Thus, there is no extended feature for other letters.

Even though the extended sections make (3) and (4) become double features, three horizontals can take (3), four dots can take (4), this will be very simple and straight forward, then it should not make big trouble.

Moreover, said basic character further comprises specially defined character. Below is the five specially defined characters of half-split input method.

- The radical "忄" can be selected in the letter (2) due to it being paired.  
(This can also avoid confusion caused by its multiple writing methods)
- The radical "扌" can be selected in the letter (6) due to it being transformed from "心".  
(This character also has multiple writing methods)
- The radical "亻" can be selected in the letter (5) due to it being transformed from "人".
- The radical "衤" can be selected in the letter (4) due to it being transformed from "衣".
- Besides, the Chinese character "心" is also a specially defined character, which can be selected in the letter (6).

As the structures are special, the definitions of these three characters are different from the general way, "本"(十, 八, 一), "州"(ㄣ, 川) and "牛"(丿, 二, 丨)

When a Chinese character is decomposed into four sections (i.e. smaller section), each section usually has become very simple. In general, there is only one basic character in a smaller section, no need to choose. However, there are also many cases where a smaller section still comprising several basic characters. Below will explain which basic character should be selected to correspond to the code.

Referring to Figure 2 and Figure 4, wherein 101 represents ten priority levels for code selection from low to high. Said encoding rule further comprises selecting said character code combination according to code selection priority of said letter feature. Preferably, said letter features in each group comprise: From low to high, the code selection priority level of letter features in said each group is represented by the number 1-10. Comparing the code selection priority level of letter feature to which each basic character in said font section belongs, wherein the corresponding character code to which one of said letter features of the highest level belongs is selected to represent said font section. For example, when a smaller section has several basic characters that can be selected for encoding, preferably select the character code that representing relatively complicated basic character.

As long as some details are ignored, it can be easily seen that the basic characters corresponding to the letter features of each letter will be gradually simplified from "four sides" to "one side" from right to left according to their position on the keyboard; gradually simplified from "four strokes" to "one stroke". Wherein, the more the number of sides or strokes, the character structure represented will also be more complicated, the intersecting character is arranged in the letter (3) which can also roughly reflect the complexity of its character structure.

Preferably, said priority level of code selection comprises: the more number of sides or/and strokes of said basic character that represented by said letter features in

each group, the higher the code selection priority level will be. Said key comprises: the position of said corresponding character code that corresponding to it on said keyboard is arranged in order from 1 to 10 according to code selection priority level of letter feature represented.

If the number in the key (0) is imagined as (10), the key name will become a sequence from (1) to (10) (equivalent to the ten levels of code selection priority of the letter features, the letter (0) represents the level 10, other letters have the same name as the level). The advantage of this is that, the answer to above question can be written in a very concise statement: compare each basic character in a smaller section, take the largest number of the corresponding letter names as the code. For example, as shown in Figure 2, the characters that belong to the letter (0) will be selected first, then will take the characters in the letter (9), then (8), (7) and so on.

In another example "塔", it is required to take a code in the smaller section "合" when encoding, but "合" comprises three letters (5, 1, 0), as (0) is equivalent to (10) which is larger, thus it is selected as the last letter of the code for "塔" (3120).

Referring to Figure 1, wherein 103 represents the respective selectable letters of different parts of "解" 102A-I, 102A-II, 102B-I and 102B-II, the dark fonts represent the ones with the highest priority level of code selection (the number represented is the largest), 104 in the figure shows the letters with the highest level are selected in each part according to the rules of code selection priority, then the code of "解" is (4762).

Since the above ten key names are arranged from right to left and from large to small on the keyboard, that means the letter on the right will have priority to be selected as the code.

In the mean time, when defining letters and arranging keys, they are arranged according to the complexity of the character features from right to left, from

complicated to simple (from 4-side to 1-side, from 4-stroke to 1-stroke), thus, "take the largest number" actually means selecting the complexity and discarding the simplicity. The half-split input method uses the common knowledge of the number and the convenient technique of "taking the largest number" to achieve this goal. The actual meaning of the following three statements are completely equivalent: select the complexity and discard the simplicity = take the right and discard the left = take the largest number.

In addition, said code selection rules comprise: when there are more than one said letter feature in highest priority level of code selection, wherein select the one in highest level, and the corresponding character code of said letter feature of basic character represented with first writing stroke order to represent said font section. In some very specific examples, the following rules are needed to determine how to select character in a "smaller section": when a "smaller section" comprises more than one "largest number" character, the one with first stroke order will be selected for encoding.

"Select the complexity and discard the simplicity" is the critical part of the half-split input method. It decomposes the characters in a completely new way. People who first contact it may have a misunderstanding: each "smaller section" in the font must be compared about its "complexity" which is wasting of time. However, if about 6,000 of the most frequently used Chinese characters are decomposed as, "one into two, two into four", they will result in many different "smaller section" fonts. Only ten kinds are comprising two different "口" (4-side) characters in so many fonts. That means only ten cases are necessary to select among the three letters (0, 9, 8). While the cases that there are two different "angle" (2-side) characters, and those which need to select among (6, 5, 4) are slightly more, are about 60 kinds. Therefore, in most cases, only need to discard the simple dot and line strokes to get the code correctly.

Moreover, the differences between a few fonts are very small and difficult to

distinguish, those will be ignored in the half-split input method. For example, "𠂇 → 王", "月 → 月", "辉 (the left section) → 光".

The ten-letter half-split input method is simple and easy to use, in order to satisfy the users who want fast and high efficiency, other keys on the keyboard can be used as the corresponding codes. Preferably, said corresponding character code comprises: there may be different letter modes, each said letter mode uses different numbers of corresponding character codes to represent different letter features in each said group of letter feature. For example, said letter mode comprises: forty letter modes comprise forty corresponding character codes, respectively use five, four or three said corresponding character codes to represent different letter features in each group.

See those ten letters of the half-split input method carefully, it may be seen indistinctly that each letter comprises multiple different elements, such as (1) comprises "a horizontal", "a vertical", "a slant" and "a dot" four features. As long as each feature is represented by a letter, while all other rules are unchanged, a new input method mode with the advantages of multiple letters and higher efficiency has been obtained. Wherein, "horizontal, vertical, slant, dot" are equivalent to the additional descriptions of said "1-stroke" character.

In order to distinguish them, the inventor named the ten-letter half-split input method as "Chinese Code" (ten-letter mode); after subdivision, the forty-letter mode is named "Goldstone Code" (forty-letter mode); but there is no subordination between them. It can be said that the ten letters of "Chinese Code" are subdivided into forty letters of "Golden Stone Code"; it can also be reversed: put several related letters in the "Goldstone Code" onto the same key to become "Chinese Code". That means can use the letter feature "1-stroke" to represent (1), or use the four letter features "a horizontal", "a vertical", "a slant" and "a dot" to represent (1).

This configuration is advantageous. For example, the ten-letter mode is suitable for mobile phones, and the forty-letter mode can be used with computer keyboards. It is

understandable that different devices have different input interfaces, which comprising the physical or virtual keyboard, and different numbers of keys can be provided on the keyboard for users to input codes.

Referring to Figure 5, the keyboard layout of "Goldstone Code" is shown. Figure 3 lists the "letter features" of each letter according to its position on the keyboard (but slightly adjusted). Figure 4 is a reference list of individual letters, letter features and the characters to which they belong.

In general, each letter will be subdivided into four new letters, each new letter subdivided by the same letter will become a group, and the same group of letters will be arranged in sequence on the original letter and the keys below it. Based on the keyboard position, for example the more common QWERTY keyboard layout, from left to right, top to bottom, respectively is:

(1) is subdivided into: (1) = "a horizontal", (Q) = "a vertical", (A) = "a slant", (Z) = "one dot" (in the description, the key names will still be used to replace the letter names).

(2) is subdivided into: (2) = "two horizontals", (W) = "two verticals", (S) = "two slants", (X) = "two dots".

(3) is subdivided into: (3) = "十" character, and there must be the "simple" stroke like a horizontal or a slant above and below it. For example, the "十" character in "王" or "壬". (E) = also the "十" character, but only require a simple horizontal stroke below, for example, the "十" character in "士" or "土". (D) = also the "十" character, which can be understood as other "十" characters other than the above groups of two letters. The following two types of characters all belong to the letter (D): 1. Only horizontal or slant stroke above it, such as the "十" character in "干" or "千". 2. The "十" character that touches "non-simple" strokes, for example, "早, 支, 孛", etc. For the same reason: the "十" characters in "至, 堂" belong to the letter (E), but not the letter (3). (C) = "乂, 𠂇" two characters.

This group of letters also has extended features, they are: (3) = "3 horizontals", (E) = "3 verticals", (D) = "3 slants", (C) = "3 dots".

It can be seen from the situations of 3 groups, the letter features of (3), (E) and (D) are "十" character. To determine which letter a character belongs to, the strokes other than "十" character must be considered, they are called "external features" (comprised in the additional description of said basic character). In the Chinese Code, the three letters (0, 9, 8) are distinguished by external feature. In the Goldstone Code, the external feature will be more important which need to be careful. It should be noted that the strokes that constitute the external feature can be used as another section for encoding.

When listing the examples that comprising external features, it is usually for the encoding part of the letter in the description, and external features are not included. In order to make the description concise, this will not be noted below. This means that in the second paragraph above, the underlined content will be omitted. That means like the underlined contents in the second paragraph above will be omitted.

(4) is subdivided into: (4) = "ㄥ", (R) = "ㄣ", (F) = "ㄚ", (V) = "ㄣ". The extended features of this group are: (R) = "4 verticals", (V) = "4 dots".

(5) is subdivided into: (5) = "人", must have external features as ---- with other strokes on both sides, to represent the "大" character. Such as "大, 火, 脊", etc. (T) = "人, 入" two characters. (G) = The upper section of "金". The difference between (G) and (T) is that there will be a related font under (G), such as "全, 食, 个", etc, (T) does not have this feature (the above three letters are all angles Pointed upwards, only the external features are different. (B) = "<", the sharp angle is to the left, such as "女, 巡, 水", etc.

(6) is subdivided into: (6) = "ㄣ, ㄣ" (angles are at the bottom). The exception is "ㄣ" in "心". (Y) = "心" (specially defined character). (H) = "ㄣ, ㄣ" (the angles are in above). (N) = "ㄣ, ㄣ, ㄣ, ㄣ", etc. Noted that the angle of the character is slightly obtuse, wherein one or both sides are slightly inclined outward (this is where it differs from (H)).

(7) is subdivided into: (7) = "匚, 凵" (the opening is to the left and right). (U) = "凵" (the opening is upward). (J) = "乚, 几" two characters. (M) = "冂" (the opening is downward), which must (and only) have the external feature of 2 or 3 horizontals inside it. Such as "月, 且, 直", etc. Noted that "丹, 用, 角, 再" are not included.

In general, the letters will be subdivided into four new letters, but the Group 7 has to borrow a key (,) from the neighbor Group 8 due to actual need. Therefore, (,) = "冂" (the same as (M) that the opening is downward). It can be understood as that, other characters with downward openings other than (M) and (J) in the Group7. For example, "巾, 丹, 同, 皿".

(8) is subdivided into: (8) = "口" character that has the external feature of only one horizontal stroke inside it. Such as "日, 曰". (I) = the character that has the external feature of a long tail. Such as "艮, 殷". (K) = "口" character that has the external feature of two horizontal strokes inside it. Such as "目, 自".

(9) is subdivided into: (9) = "口" character that has two strokes of straight line, oblique line or curve from up to down inside it. Such as "𠂇, 曲, 四". (O) = "口" character that has one stroke of straight line or oblique line from up to down inside it. For example "中, 史, 巴". (L) = It can be understood as that, all characters that are excluded in other letters of Group 9. Such as "因, 国, 田, 由, 甲", etc. (.) = "口" character that has the external feature of a "十" inside it, but the central line must go through the top and bottom. For example "申, 更, 电" (noted that "田, 由, 甲" are not included).

(0) is subdivided into: (0) = "口" character that does not touch other strokes. For example, "合, 沿, 或". The exception is the "口" character under "言". (P) = "口" character under "言". The character also has a long tail, such as "巳, 尸, 追". (;) = "口" character that has other strokes touching it below. For example "足, 兄, 京, 豆". (/) = "口" character that has other strokes touching it above. For example "古, 石, 名, 各".

There are very few "口" character fonts with the strokes touching both above and below, their letters should be defined by the structure that was later decomposed. For example "克", because it will be first decomposed into "十, 兄" two sections, so its "口" character will take the letter ";" according to "兄".

Based on the different writing habits, whether it is touching or not, is just an approximate description, and the attached examples should prevail.

For the other four specially defined characters, it is easier to remember by using their Chinese letter names (the key names are listed in parentheses). They are: "丩" = "心" (Y), "亻" = "人" (T), "衤" = "衣" (V), "𠂇" = "华" (W) .

After each letter is "subdivided", the previous number will still be used when "selecting the complicity and discarding the simplicity". That means each letter in the same group has the same "number". Thus, the selection will be based on the stroke order. In this way, the characters selected for encoding are completely consistent with the "Chinese Code".

The two main letter modes of the half-split input method have all been introduced until here. Other modes are auxiliary. For example, since forty letters can be arranged in groups on ten keys to get a different letter mode, can also create a twenty-six letter input method mode to cater to users who are accustomed to twenty-six English letters.

Figure 7 shows the keyboard layout of this letter mode ---"26-Code". To combine the letters of the upper row of numeric keys with the letters of the lower keys, and the letters on the right symbol keys are moved to the left and combine into the adjacent keys. Only need to operate according to the letter names (a key will comprise multiple letter names) of "Goldstone Code" on each key in Figure 7 (in fact, the letters are not reduced, only the keys are reduced). However, it should be noted that the two letter features of (P), the "口" character under "言" and with a long tail, will be combined into different keys (P) and (L).

Although the input method modes with different numbers of letters can be optionally created, It cannot guarantee that each mode can meet the requirements of easy learning and high efficiency. The half-split input method does not combine or separate different letters at random. On the contrary, after numerous times of matching and repeated corrections, the letters are still related to each other when they are combined together, so that they are not messy or out of place. When they are "separated", they have their own distinct features, so it is not easy to confuse them. Finally, each mode has two advantages: low character selection rate and easy to master.

The multi-letter mode will definitely have more letter features. If different letter features should be considered in different modes when decomposing the fonts, it cannot maintain their simplicity and consistency. However, the decomposition method of the half-split input method only considers the nine simplest basic characters, and does not involve letter features. Therefore, the only difference between a multi-letter mode user and a ten-letter mode user is: he only need to judge which one of the several letters of the group it belongs to from the character direction of the basic characters, the difference between dots and lines, the difference in stroke angle, and the external features (other rules are unchanged).

In the examples of Chinese Code, Goldstone Code, 26-Code mentioned above, when said input code exactly matches the records of said Chinese character encoding table, the matched Chinese character can be input into the device. Alternatively, the input method may be set to receive an input code combination that matching (at least) a part of the records of said Chinese character encoding table, and the matched Chinese character can be input into the device. For example, display all Chinese characters that at least partially match the records in a text index mode, allowing the user to select and input the required Chinese character into the device.

Preferably, said key representing each input code comprises: one said key may correspond to two corresponding codes. Selecting/ pressing the said key is equivalent to inputting said two corresponding character codes.

Another letter mode can be called "One Code". In this example, the letters of the Chinese Code can be combined into a series of new letters two by two (more than one hundred). Such as (12, 13, ..., 91, 92), etc. Take "昇" as an example to demonstrate how to use "One Code" mode to input text. The Chinese code of "昇" is (0612), the first part of the code is (06), so press the key of (06), then the character selection table will show the character list according to the second part of the code, find "昇" in the (12) position, then complete the input. Although there are many letters and characters can be selected, they are all arranged in sequence. As long as taking proper training, the texts can be input quickly. In above example, if the code (06) has too many characters, then also press the key (12) to input the complete code. Take a slight modification of this mode, it can become a systematic text indexing method similar to traditional dictionaries with similar operations.

In addition to the different letter modes, the half-split input method has two other encoding modes (equivalent to the different modes of said decomposition rules) to expand the scope of application of the half-split input method. Preferably, said half-element character comprises the thirty-four most frequently occurring radicals or component characters in Chinese characters, and the characters which are similar to them.

The inventor considers the mode in this example to be an advanced mode, hereinafter referred to as "half-element method". Preferably, this mode can make the character selection rate (the probability of needing to select the Chinese character after inputting the codes of a Chinese character) to a minimum without increasing the number of codes (the average of the number of letters included in the code of each Chinese character). It can be set as part of the basic mode of the half-split input method.

The input efficiency of Chinese Code is sufficient to surpass other input methods with similar numbers of letters even it does not use the half-element method. In consideration of allowing beginners to master and apply as soon as possible, the "half-element method" is set as advanced skills, the user can decide whether to learn or not.

The critical part of the half-element method is thirty-four half-element characters. Referring to Figure 6, each half-element character is arranged according to the position of the letter on the keyboard. If they are compared with the letters of Goldstone Code in Figure 5, it can be seen that they are very similar except partially replaced by their own component characters.

Half-element characters can only be regarded as "pre-half-element", because they must be in the correct positions, otherwise they cannot become the real "half-element". The correct positions can be simply understood as the most frequently occurring positions of these characters (one of the upper, lower, left, and right part of a Chinese character). Below is the strict descriptions.

For "讠, 犭, 艹, 艹, 辶, 心", etc half-element characters, as they will only appear in the upper or lower half (approximately equivalent to one of the sections after decomposing into two sections) of a Chinese character, so do not need to remember. The rest of the half-element characters, except "鳥", must be located in the left half. Just remember that "鳥" is the only right half, it must be in the right half.

After understanding how to correctly identify the half-element, then start to understand how they are coded. The encoding rules may comprise using one character code to represent said half-element character, and using a combination of three character codes to represent the remaining part of said Chinese character other than the half-element character. For example, when a Chinese character comprises the half-element, the half-element itself takes only one code, while the rest of the Chinese character (hereinafter referred to as "the other half") takes three codes.

The other half will be decomposed (two are divided into four) into two sections (i.e. smaller section), the first section can be understood as the "first part" (the part with the first stroke order), and the rest are called "second part" (the part with the latter stroke order). The rules of taking three codes for the other half are: "first part" will take two codes at first, and "second part" will take one code. If the "first part"

only has one code (only one basic character that cannot be divided), then the "second part" can take two codes.

If the "first part" is also a "half-element", the situation will be completely reversed. "First part" only takes one code, while "second part" can take two codes at first. If the "second part" has only one code, the "first part" that is also the "half-element" can still take two codes. Taking two codes means "one is divided into two, each taking one code". This is the standard encoding mode for half-split input method.

Specifically, the half-element method defines the most frequently occurring character structures in Chinese character structures (thirty-four half-element characters), and then minimizes the "seats" they occupy in encoding. "Half-element" cedes one of the two "seats" to "the other half"; and the "first part" belongs to the half-element character will also give the "second part" the opportunity to take two codes. In this way, the different sections of the font are easier to be selected for encoding, to avoid the same code situation, thereby reducing the character selection rate.

Alternatively, the encoding rules may comprise representing the said half-element character with a combination of two character codes, and representing the remaining section of said Chinese character other than the half-element character with a combination of three character codes. When using the half-element method in the "26-Code", wherein ten of the half-element "艹, 王, 糸, 火, 巾, 目, 酉, 車, 言, 石", will not be the same as the ordinary half-element that taking only one code, they still have to take two codes (so the codes of some Chinese characters in the "26-Code" are five letters). This can make up for the shortages of this mode due the small number of letters, and the character selection rate being slightly higher than the "Goldstone Code".

The method of taking three codes for "the other half" can have another simpler form. The rules are as follows: If both "first part" and "second part" have two codes, then eliminate the one with the smallest number.

The half-element method can be regarded as the most suitable for users who seek speed and efficiency. However, if the users do not mind having a proper amount of character selection, can also simplify the encoding method of the half-split input method starting from the opposite direction. For example, the encoding rules may comprise limiting the said character code combination to have one to two character codes. In this example, when typing each Chinese character, only need to press at most two letters. This mode can be understood as a simple mode, the 2-Code mode, its codes are taken from: One divided into two, take one code each.

Preferably, the following explanations can be used to distinguish the difference among the three encoding modes: General (without half-element method) = two plus two, half-element method = one plus three or three plus one, 2-Code mode = one plus one.

In addition, simplified Chinese characters can also use the half-split input method. Figure 8 - 9 shows the keyboard layout of the half-split input method of simplified characters. Below it is the half-element of simplified Chinese characters.

Referring to Figures 5 - 6 at the same time, the most obvious difference between the traditional and simplified "half-split" is the letter names (Chinese letter names are convenient for expression, the half-split input method can still work as usual if only the key names are used). As in the simplified Chinese characters, the three original letters "车(車), 鸟(鳥), 华(華)" must be moved to other keys (are no longer the letters) due to the change of the character, the original positions will be replaced by "申, 艮, 卅".

In addition, the radical writings of "言, 金" have changed greatly in simplified Chinese characters. Therefore, the special definition method is adopted to deal with them, so as to avoid too big difference between the results of code selection of traditional and simplified Chinese characters.

讠, the radical writing of "言", the whole character takes the letter "言" (P) / Chinese Code (0).

车, the radical writing of "金", the whole character takes the letter "金" (G) / Chinese Code (5).

In terms of the "half-element method", for the two half-element that are "migrated", since the new feature of "车", "厶" already has another half-element on the key to which it belongs, it must be eliminated from the half-element family. "鸟" is the right half-element, although it is on the same key with "马", it can still retain the position of half-element.

The input method of the invention can further provide the operation skills of the following auxiliary functions. Such as inputting words (terms), inputting the traditional Chinese characters to output the simplified Chinese characters, etc.

The rules for inputting words comprise: 2-character word, two codes are taken from two characters. 3-character word, two codes are taken from the first character, the remaining two characters take one code each. 4-character word, one code is taken from each character. The word with more than four characters, one code is taken from each of the first three characters and one code is taken from the last character.

Other functions can help users who forget the writings of the Chinese characters, which is adding "additional key" (not the keys used for the letters of half-split input method, such as ], [, \, =, -, etc.) or "additional letter" (the letters of half-split input method) in front of the codes. Since the latter is the letter key, when used in conjunction with a Chinese character that has only three letters or less, the "additional letter" must be repeated so that the total number of codes reaches five (when over half-element is used, it should be six) letters. To avoid confusion with the encoding of ordinary Chinese characters, the following examples illustrate:

Input homophone: Add the additional key (-) or additional letter (0) in front of the code of "美" to input the homophone "每". (The symbols ( ) will be used to mark additional keys and additional letters below, and will not be further described.)

Input the character with similar pronunciation: Add (=) or (9) in front of the code of "美" to input the Chinese character with similar pronunciation "微".

Use the first Chinese character of a word to input: Add ([) or (8) in front of the code of "美", the first Chinese character of the word "美观" to input the second character of the word "观".

Use the last Chinese character of a word to input: Add (]) or (7) in front of the code of "观", the second Chinese character of the word "美观" to input the first character of the word "美".

Input the traditional character to output the simplified character: Add (') or (6) in front of the code of "華" to output its simplified character "华".

Input the simplified character to output the traditional character: Add (\) or (5) in front of the code of "华" to output its traditional character "華".

When inputting words, in order to avoid the word code is the same with the general single Chinese character code, thus increase the character selection rate, can also add additional letter in front of the word code. Comprising: the word with four or more characters can use (4), 3-character word can use (3), 2-character word can use (2).

Since when the Goldstone Code is used in conjunction with the half-element method, the character selection rate is very low, another code can be set for each character that needs to be selected to avoid the operation of selecting characters.

Substitution Code: Add (S) in front of the code of "居", or add (2) in front of the code of 2-Code mode of "居" to get "居" without character selection. (The second method can ignore the requirement that the total number of codes with "additional letter" is no less than five, but there are additional rules for some characters as needed.)

The above functions are only the arrangement of encoding, there are no additional requirements for input method procedures. In addition, the half-split input method further provides some usage modes that require to work with the input method program. For example: when Caps Lock is enabled, can use the word codes to input the word. In this way, not only the character selection rate will not be affected, but also no need to add the additional letter.

It can be seen that although the half-split input method is named "split", its true intention is exactly the opposite. Advantageously, the half-split input method can find the best balance of "easy to learn" and "efficiency", and comprises both "multi-letter" and "less-letter" modes and their advantages. Finally, let the Chinese character input method unify in a simple and effective encoding method.

Therefore, after describing several examples, the technical personnel in this field may realize that different modifications, additional structures, and equivalents can be used without departing from the essence of the invention. Accordingly, the above description shall not be deemed as limiting the scope of the invention as determined by the claims of the application.

## Claims

1. A Chinese character input method, wherein:

Firstly, encode Chinese characters and establish a Chinese character encoding table, the encoding method comprises following steps:

Decompose the Chinese characters or words needing to be input into one or more font section(s) based on the decomposition rules, wherein each of said font sections comprises: at least one basic character; and

Based on the code selection rules, respectively select a corresponding character code to represent each of said font sections according to said basic characters contained therein; and

Based on the encoding rules, combine the selected corresponding character codes into character code combinations, to represent said Chinese characters or words needing to be input;

Wherein, said font section comprises: based on said decomposition rules, said Chinese characters needing to be input or several Chinese characters of said words needing to be input that do not need to decompose or cannot decompose; Wherein, said word comprises: respectively combined by two or more Chinese characters;

Then, receive the input code combination that input by keyboard, match the input code combination with said character code combination that representing each Chinese character or word in the Chinese character encoding table; and

Finally, display the Chinese characters or words that represented by said input code combination;

wherein, for the keys/options representing different corresponding character codes, their configurations on said keyboard will be arranged according to code selection priority level of corresponding letter feature of said corresponding character codes;

wherein, said basic character comprises: 4-side character, 3-side character, 2-side character and 1-side character that defined by the number of sides, 1-stroke character, 2-stroke character, 3-stroke character and 4 stroke character that defined by the number of strokes, and intersecting character that only comprising two intersecting lines; wherein the different rules of stroke structure of the same

said basic character or/and other than said basic character, represent as additional descriptions of said basic character;

wherein, said 4-side character comprises: the character roughly displayed as a square "□"; said 3-side character comprises: the character that any one side of the square "□" is missing; said 2-side character comprises: the character that any one side of the 3-side character is missing, comprising the character roughly having one angle of square or rhombus, the character having an acute angle or an hook shape, wherein said square angle, rhombus angle, acute angle and hook shape represent as additional descriptions of said 2-side character;

wherein, said 1-stroke character, 2-stroke character, 3-stroke character and 4-stroke character comprise: the horizontal, vertical, slant and/or dot stroke, having 1-4 stroke(s) that do not intersect with each other, link together, and do not intersect with other strokes at all or all intersect with one or more other strokes;

wherein said horizontal, vertical, slant and dot represent as additional descriptions of said 1-4 stroke characters; wherein a horizontal, a vertical and a slant stroke of said 1-stroke character are equivalent to said 1-side character;

wherein, a said basic character or a basic character with additional description is represented by a letter feature; wherein the same basic character with different descriptions are represented by different letter features;

wherein, said letter feature is classified into ten groups, the basic character that represented by each said letter feature in the same group having the same number of sides or/and strokes;

wherein, said letter features in each group will be represented by one or more said corresponding character code(s) in different modes; wherein, in comparison with the mode having relatively less said corresponding character codes, the same basic character has relatively more additional descriptions in the mode having relatively more said corresponding character codes, so that increase the number of letter features in the letter feature group to which it belongs and meet the needs of this mode;

wherein said letter features in each group match with one said code selection priority level from low to high marks; wherein, the more number of sides or/and strokes of said basic character that represented by said letter features in each group, the higher the code selection priority level will be;

wherein, said code selection rules comprise: comparing the code selection priority level of letter feature to which each basic character in said font section belongs, wherein the corresponding character code to which one of said letter features of the highest level belongs is selected to represent said font section;

wherein, said decomposition rules of Chinese character are classified into: 2-code mode, general mode, half-element mode and over half-element mode; said modes of decomposition rules further comprise word mode; wherein said modes of decomposition rules comprise: when each of said font sections comprises more than one basic character, and needs to be decomposed into two relatively smaller font sections according to the regulations of said mode of decomposition rules, then decompose said font section into two relatively smaller font sections based on decomposition method; and repeat until finishing the specified steps of said mode of decomposition rules; when said font section needing to be decomposed has only one basic character that cannot be decomposed any more, then ignore said step;

Wherein said decomposition method comprises: Preference is given to decompose the font section needing to be decomposed into two font sections that each containing a Chinese character, a radical or a common character, or/and decompose into two font sections with similar size or square shape, or/and decompose into two font sections that less strokes contacting or overlapping with each other in the original font section;

Wherein, the decomposition steps specified in said 2-code mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method;

Wherein, the decomposition steps specified in said general mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method, then further decompose each of acquired font sections into two smaller font sections respectively based on said decomposition method;

Wherein, the decomposition steps specified in said half-element mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method; when one of two acquired font sections is half-element character and said half-element character locates in specified

position in the Chinese character needing to be input, then further decompose another font section outside said half-element character into no more than three font sections based on specified decomposition steps of the section outside half-element character;

Wherein, the decomposition steps specified in said over half-element mode comprise: decompose the Chinese character needing to be input into two font sections based on said decomposition method; when one of two acquired font sections is over half-element character and said over half-element character locates in specified position in the Chinese character needing to be input, then further decompose said over half-element character into two smaller font sections based on said decomposition method; and further decompose another font section outside said over half-element character into no more than three smaller font sections based on specified decomposition steps of the section outside said half-element character;

Wherein, the decomposition steps specified in said section outside half-element character comprise: decompose said another font section outside half-element character or over half-element character into two smaller font sections based on said decomposition method; in two acquired smaller font sections, when the one with first writing stroke order is also a half-element character or only comprises a basic character, then further decompose the one with latter writing stroke order into two smaller font sections based on said decomposition method; and when said one with first writing stroke order is not a half-element character or said one with latter writing stroke order only comprises a basic character, then further decompose said one with first writing stroke order into two smaller font sections based on said decomposition method;

Wherein, the decomposition steps specified in said section outside half-element character further comprise: further decompose said another font section outside half-element character or over half-element character into two smaller font sections based on said decomposition method; and further decompose each of acquired font sections into two smaller font sections respectively based on said decomposition method; and when more than three font sections acquired from above steps, then compare the code selection priority level corresponding to corresponding character code for each of the font sections acquired from above

steps, remove one said font section which has lowest priority level of code selection and the writing stroke order of its corresponding basic character is latter, will not be used as said font section for encoding.

2. The Chinese character input method of Claim 1, wherein said encoding rules comprise: each of said corresponding character codes in said character code combination is arranged according to writing stroke order of its corresponding basic character.

3. The Chinese character input method of Claim 1, wherein said half-element character and over half-element character are the most frequently occurring radicals or component characters in Chinese characters, and the characters which are similar to them;

Wherein said specified position of half-element character and over half-element character is where said half-element character most frequently occurring in Chinese characters.

4. The Chinese character input method of Claim 1, wherein the decomposition steps specified in said word mode comprise:

When it is a 2-character word needing to be input, decompose the first and second character of said 2-character word into two font sections respectively based on said decomposition method;

When it is a 3-character word needing to be input, decompose the first character of said 3-character word into two font sections based on said decomposition method, and use the second and third character of said 3-character word to act as a font section respectively without decomposition;

When the word needing to be input comprises more than three characters, use the first, second, third and last character of said word to act as a font section respectively without decomposition, the other characters will not be used as said font section for encoding.

5. The Chinese character input method of Claim 1, wherein said Chinese characters displayed matching with said input code combination comprise: When need to display several Chinese characters in said record that partially matching with said input code combination, the arrangement order of said each Chinese character is displayed from low to high according to code selection priority level for the letter

features represented by the corresponding character codes in character code combination encoded in said each Chinese character.