(15)

PATENT SPECIFICATION (11) 1 550 144

1 550 144

(21) Application No. 41617/76 (22) Filed 7 Oct. 1976 (19)

(31) Convention Application No. 690621 (32) Filed 27 May 1976 in

(33) United States of America (US)

(44) Complete Specification Published 8 Aug. 1979

(51) INT. CL.[2] G09C 1/00

(52) Index at Acceptance
G5X 30



(54) METHODS FOR ENCODING BINARY WORDS

(71) I, CHAN-HUE YEH, a citizen of the United States of America, of c/o 1031 E. East Duane Avenue, Sunnyvale, California 94086, United States of America, do hereby declare the invention, for which I pray that a patent may be granted to me, and the method by which it is to be performed, to be particularly described in and by the following statement:–

This invention relates to methods for encoding binary words.

This invention is related to my prior U.S. Patent 3,820,644 issued 28 June 1974 and entitled "System for the Electronic Data Processing of Chinese Characters", the disclosure of which is incorporated herein by reference. In my aforesaid prior U.S. patent there is described a method by which each of the approximately 10,000 Chinese characters can be assigned a unique 14-bit binary code so that each character is uniquely identified. For example the Chinese character 北 (Pei or North) is assigned a unique decimal numer 10165 which has an equivalent binary representation 01000100000101. A keyboard is so designed that each character can be easily accessed electronically by pressing not more than three keys on the keyboard at speeds far in excess of that possible by human encoding. It is, however, sometimes desirable that Chinese characters by both human-readable and machine-readable and it will be appreciated that a standard Chinese typewriter is not adapted to type both human readable characters and the somewhat cumbersome 14-bit binary code which is machine readable. It is, therefore, desirable that the machine readable 14-bit binary code be further reduced to a line code which is also machine readable. Line codes for representing the Roman alphabet are known to the art, for example in U.S. Patent 3,657,522 entitled "Record for Machine Scanning" issued 18 April 1972 to Wildhaber, in which each letter of the alphabet is assigned a number from 1 to 26 which is then transcribed in binary form so that, for example, the second letter of the alphabet B can be written as 00010 or as merely a bar in the fourth position from the left or start. For convenience a short bar is provided preceding each coding to indicate the start of a code, thus A can be represented by '. . . .1, B can be represented by '. . .1., C can be represented by '. : .11, and so on for all 26 characters of the alphabet. It will be appreciated that even 26 characters requires a relatively large and clumsy binary layout containing a marker and 5 binary bar positions. The difficulties encountered when encoding 10,000 or more characters are immeasurably greater, and as has been indicated in my prior patent a 14-bit binary notation is the simplest which can be used to uniquely identify all 10,000 or so Chinese characters. Even with modern scanning equipment which can read or scan lines 0.01" wide and 0.2" high the space required for a 15 bar code (a market bit plus 14 information bits) is too great for the bar code to be incorporated on a character font of a Chinese typewriter.

It will be appreciated that although this invention is particularly suitable for encoding identity information for Chinese characters, it is not restricted thereto. Other forms of information which can be represented in a binary notation can also be encoded according to the principles of the present invention. While the invention is particularly suitable for encoding 14-bit binary representtations the invention is equally applicable to any longer or shorter even numbered binary representations.

According to the present invention, there is provided a method for encoding a binary word consisting of an even number of binary bits, which comprises:

(a) assigning a line of a selected length to each 0-bit and a line of a different

selected length to each 1-bit of said binary word;

(b) establishing a first line code for said binary word by disposing said lines in a side by side parallel relationship along a common base;

(c) dividing said first line code into equally sized first and second groups of lines;

(d) placing said first and second groups one above the other and inverting one of said groups so as to place the common base of the first said group adjacent the common base of the second said group; and

(e) aligning the lines of said first and second groups in abutting end to end relationship whereby to form a second line code which is uniquely representative of said binary word.

An embodiment of the invention will now be described by way of example, with reference to the accompanying drawings, in which:

Figure 1 is a perspective view of a Chinese typewriter provided with a plurality of lead font character blocks;

Figure 2 is a perspective view of one of the lead font character blocks;

Figure 3 is a sketch of a machine-readable 14-line code representing a 14-bit binary representation;

Figure 4 is a sketch of 14-line code of Figure 3 broken into two 7-line segments;

Figure 5 is a sketch of the 14-line code of Figure 4 in which the bottom 7-line segment is inverted;

Figure 6 is a sketch showing the segments of Figure 5 joined to form a new 7-line code;

Figure 7 is a sketch shvwing the code of Figure 6 with a lead or start line;

Figure 8a is a sketch showing the 4 possible line lengths, and positions for the code of Figure 6; and

Figure 8b is a sketch showing a further line code derived from code of Figure 7.

Turning firstly to Figure 1 there is shown a Chinese typewriter 1 having a base 2, a slidable carriage 3 and roller 4 around which the record medium 5 such as a sheet of paper is wound in conventional manner. In place of the conventional keys in a standard western-style typewriter, a plurality of rows of lead font character blocks 6 are provided each adapted to the lifted out of its storage position to a printing position by means of a hammer mechanism 7 on movement of selector 8 and actuator lever 9. In a standard Chinese typewriter the character fonts are similar in size and shape to the font 6 shown in Figure 2 but are embossed with the appropriate human-readable character 10 only. The font 6 is also provided with an 8-line machine-readable code 11 which is generated according to the principles described in more detail hereinafter.

As described in my prior U.S. Patent 3,820,644 each of the approximately 10,000 Chinese characters can be assigned a unique 14-bit binary code which serves to uniquely identify each character. For example the character 10 ("Pei" or North) is assigned the number 10165 which can be defined in an equivalent binary form as 01000100000101. As machines can be programmed to read binary codes either on the basis of go-no-go or by scanning the length of a line, a machine-readable code in which both the 0 and 1 are represented by a line can be developed. Thus a short line represents 0 and a long line (twice the length of the short line) represents 1, and the 14-bit binary code for Pei can be represented as shown in Figure 3, the lines being disposed in a side by side parallel relationship along a common base. This 14-line code is still too long to be incorporated onto a font 6, which is only of the order of 0.25" wide. The next step, therefore, is to break the 14-line code into two 7-line groups and place one group vertically above the other as shown in Figure 4. In order to prevent undue or unnecessary super-imposition the bottom group is inverted about a line parallel to its base, as shown in Figure 5. The two segments are then joined as shown in Figure 6 to provide a 7-line code which uniquely defines the 14-bit binary representation of the decimal coding of each Chinese character. It will be appreciated that there are only four possible notations for the derived code lines, which vary in length from 2–4 lengths inclusive, that is 4 lengths (2–2), 3 lengths (2–1, 1–2), and 2 lengths (1–1) — the numbers in brackets referring to the line references in Figures 6 and 8a always reading from the top downwardly. Surprisingly, it has been found that these four possible lengths and relative positions are sufficient to uniquely define the 14-bit binary codings necessary to define all 10,000 or so Chinese characters.

Optical readers are, of course, known in the art and commercially available readers are capable of reading or scanning lines which are 0.01" wide and 0.2" high and which are spaced by 0.01". Thus the 7-line code described above can be compressed into a space of 0.14" and hence can be embossed on the font of Figure 2, so that both machine and humanly readable characters can be provided on a single font.

In a preferred embodiment of the present invention, a lead line is added to the 7-line code, as shown in Figure 7, in order to indicate to the reaching machine the start of the line code and to reset the machine status. In this embodiment it will be appreciated that the width of the machine-readable code will be about 0.16"

It will also be appreciated that Chinese may either by read in horizontal style, from

left to right or in vertical style from the top down and from right to left. The human-readable and machine-readable code may therefore be embossed on the front one above the other, as shown in Figure 2, or side by side (not shown).

In a further embodiment of this invfntion the 7 or 8-line code can be modified to be compatible with alternative optical readers using different logic designs and which are also commerically available. Thus the four lines described hereinabove with reference to Figure 8a can be represented by uniform length bars of different width as desired. As shown in Figure 8b the shortest line (1–1) is represented by a bar of zero width 81, line (2–1) is represented by a bar of say 0.01" width 82, line (1–2) is represented by a bar of 0.02" width 83, and line (2–2) is represented by a var of 0.03" width 84, so that the 8-line code of Figure 6 can be written in an equivalent form as shown in Figure 8b. Other modifications of the final readout form code will suggest themselves to those skilled in the art, without departing from the scope of this invention as defined in the apprended claims. As previously indicated, although specific reference has been made to the encoding of information to uniquely define Chinese characters, the principles of the encoding system can be applied to many other forms of information which can be represented by an even number of binary bits and a line code can be developed with any number of lines depending upon the complexity of the information which is codified.

Attention is drawn to my patent application no. 39033/78 (Serial No 1550145) divided out from this application.

WHAT I CLAIM IS:-

1. A method for encoding a binary word consisting of an even number of binary bits, which comprises:

(a) assigning a line of a selected length to each 0-bit and a line of a different selected length to each 1-bit of said binary word;

(b) establishing a first line code for said binary word by disposing said lines in a side by side parallel relationship along a common base;

(c) dividing said first line code into equally sized first and second groups of lines;

(d) placing said first and second group; and

(e) aligning the lines of said first and second groups in abutting end to end relationship whereby to form a second line code which is uniquely representative of said binary word.

2. A method as claimed in claim 1 wherein said step of inverting said group about aline parallel to said common base thereof.

3. A method as claimed in claim 1 or claim 2 wherein said binary word has 14 bits and represents a Chinese character.

4. A method as claimed in any one of claims 1 to 3 wherein one said selected length is twice the other selected length.

5. A method as claimed in any one of claims 1 to 4 including the further step of assigning equal length lines of a different selected width (which may include zero width) to each combined line of different length in said second line code, whereby to form a third line code which can be optically scanned.

6. A method as claimed in claim 5 wherein the width of said lines of said third line code vary incrementally from zero to three line widths.

7. A method as claimed in any preceding claim further including inscribing a humanly readable Chinese character and a said second or third line code representative of said character onto a single character font for a Chinese typewriter, so that humanly readable and machine readable charagers can be typed simultaneously.

8. A method as claimed in claim 1 and substantially as described herein with reference to the accompanying drawings.

MICHAEL BURNSIDE & PARTNERS,
Chartered Patent Agents,
Hancock House,
87 Vincent Square,
London SW1P 2PH.
Agents for the Applicant

Printed for Her Majesty's Stationery Office, by Croydon Printing Company Limited, Croydon, Surrey, 1979.
Published by The Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.

left to right or in vertical style from the top down and from right to left. The human-readable and machine-readable code may therefore be embossed on the front one above the other, as shown in Figure 2. or side by side (not shown).

In a further embodiment of this invfntion the 7 or 8-line code can be modified to be compatible with alternative optical readers using different logic designs and which are also commerically available. Thus the four lines described hereinabove with reference to Figure 8a can be represented by uniform length bars of different width as desired. As shown in Figure 8b the shortest line (1–1) is represented by a bar of zero width 81, line (2–1) is represented by a bar of say 0.01" width 82, line (1–2) is represented by a bar of 0.02" width 83, and line (2–2) is represented by a var of 0.03" width 84, so that the 8-line code of Figure 6 can be written in an equivalent form as shown in Figure 8b. Other modifications of the final readout form code will suggest themselves to those skilled in the art, without departing from the scope of this invention as defined in the apprended claims. As previously indicated, although specific reference has been made to the encoding of information to uniquely define Chinese characters, the principles of the encoding system can be applied to many other forms of information which can be represented by an even number of binary bits and a line code can be developed with any number of lines depending upon the complexity of the information which is codified.

Attention is drawn to my patent application no. 39033/78 (Serial No 1550145) divided out from this application.

WHAT I CLAIM IS:-

1. A method for encoding a binary word consisting of an even number of binary bits, which comprises:

(a) assigning a line of a selected length to each 0-bit and a line of a different selected length to each 1-bit of said binary word;

(b) establishing a first line code for said binary word by disposing said lines in a side by side parallel relationship along a common base;

(c) dividing said first line code into equally sized first and second groups of lines;

(d) placing said first and second group; and

(e) aligning the lines of said first and second groups in abutting end to end relationship whereby to form a second line code which is uniquely representative of said binary word.

2. A method as claimed in claim 1 wherein said step of inverting said group about aline parallel to said common base thereof.

3. A method as claimed in claim 1 or claim 2 wherein said binary word has 14 bits and represents a Chinese character.

4. A method as claimed in any one of claims 1 to 3 wherein one said selected length is twice the other selected length.

5. A method as claimed in any one of claims 1 to 4 including the further step of assigning equal length lines of a different selected width (which may include zero width) to each combined line of different length in said second line code, whereby to form a third line code which can be optically scanned.

6. A method as claimed in claim 5 wherein the width of said lines of said third line code vary incrementally from zero to three line widths.

7. A method as claimed in any preceding claim further including inscribing a humanly readable Chinese character and a said second or third line code representative of said character onto a single character font for a Chinese typewriter, so that humanly readable and machine readable charagers can be typed simultaneously.

8. A method as claimed in claim 1 and substantially as described herein with reference to the accompanying drawings.

MICHAEL BURNSIDE & PARTNERS,
Chartered Patent Agents,
Hancock House,
87 Vincent Square,
London SW1P 2PH.
Agents for the Applicant

Printed for Her Majesty's Stationery Office, by Croydon Printing Company Limited, Croydon, Surrey, 1979.
Published by The Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.