

# 非确定有穷自动机在盲文转码中的应用

张居晓

(河海大学计算机与信息学院 南京 211100) (南京特殊教育师范学院信息科学学院 南京 210038)

**摘要** 对盲人使用的计算机交互技术进行研究是很有意义的工作。盲文字库缺少国际标准和不同公司的盲文字库不兼容造成了很多难题。将盲文用汉字点位编码表示,能摆脱盲文字库的束缚。用非确定有穷自动机描述盲文与汉字点位编码的转换过程,再用逆序拆分子集法对其确定化。经测试系统的转码正确率达到100%,从而实现盲文与机器无关,使得盲人使用计算机更方便。

**关键词** 非确定有穷自动机,确定化,逆序拆分子集法,盲文转码

**中图分类号** TP391.1 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.01.050

## Application of Nondeterministic Finite Automata in Braille Transcoding

ZHANG Ju-xiao

(College of Computer and Information, Hohai University, Nanjing 211100, China)

(College of Information Science, Nanjing Normal University of Special Education, Nanjing 210038, China)

**Abstract** It is of great significance to conduct research on computer interactive technology for the blind. However, due to lack of international standard, braille fonts of different companies are incompatible with each other, and this has caused a lot of problems. If braille can be presented through Chinese characters point-location encoding, it could be freed from restrictions of braille fonts. The paper presented the transcoding process from braille to Chinese character point-location encoding through nondeterministic finite automata. Then, it is verified through the method of Reverse Order-Splitting-Sets method. It is tested that the transcoding accuracy rate reaches 100%. In this way, braille is computer independence, so that the blind could use computers in a more convenient manner.

**Keywords** Nondeterministic finite automata, Deterministic, Reverse order-splitting-sets method (ROSS), Braille transcoding

## 1 引言

信息技术已经融入每个人的工作、学习和生活中,使用计算机进行信息处理是基本能力,但是目前普通计算机中信息处理的人机交互是针对普通人设计的,较少考虑到视觉障碍人士的应用需求<sup>[1]</sup>。据统计,2010年末我国约有视力残疾人1263万<sup>[2]</sup>,世界卫生组织估计全世界有盲人4000万到4500万,低视力的数量是盲人的3倍,约1.4亿人。日益增长的残疾人口和他们对信息迫切的需求是目前信息无障碍建设面临的严峻挑战<sup>[3]</sup>。

### 1.1 研究现状

盲人使用普通计算机的人机交互包含输入与输出两个方面,所处理的信息是除了普通人使用的文字(包括中文、英文、各类符号等),还多了一种文字(盲文点字符)。输入阶段是指通过键盘、鼠标、麦克风等输入设备将文字或命令输入计算机;输出阶段则是指将文字或其他信息以视觉或语音方式输出给用户。由于语音识别技术和语音合成技术已经日益成熟,采用文字与语音之间的相互转换技术进行输入输出都已成为现实,盲人也主要使用这样的方法,尤其在智能手机、平板电脑上广泛使用。

为了解决盲人使用计算机的人机交互问题,国内外研究者进行了大量研究,开发了大量的系统或者软件。输入设备或系统包括键盘(针对盲人的输入法)、麦克风(语音识别技术:语音到文字的输入技术)和扫描仪(OCR技术:纸质文字转换为电子文档技术);输出设备或系统包括点显器(文字到盲文输出设备)、盲文刻印机(文字到纸质盲文输出设备)和读屏软件(文语转换技术:文字到语音的输出技术)等。这些技术中专门针对盲人设计的包括:盲文字符OCR是将纸质盲文图片识别成盲文字符<sup>[4-7]</sup>,再转换成汉字<sup>[8-13]</sup>;使用键盘可以输入汉字,再把汉字转换为盲文<sup>[14,15]</sup>;也可以使用键盘输入直接盲文<sup>[16-19]</sup>,用点显器输出盲文,或者将盲文转换成汉字输出;点显器能够将计算机中的文字(汉字和盲文)转换为盲文输出<sup>[20-22]</sup>。目前国内相对较成熟的系统是阳光软件<sup>[23]</sup>,它具有输入与显示汉字和盲文、盲文与汉字互转等功能。

### 1.2 问题与不足

盲人使用计算机的交互技术研究是一个很有意义的工作。除去普通人也能用的技术(语音识别、文语转换等)外,其他技术均要涉及一个基础工作,即盲文字库。有了盲文字库的计算机才能输入、存储、显示盲文。目前国际上Unicode的编码的2800-28FF区域是braille点字,但Unicode的盲文是8

点字盲文(如 $\begin{matrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{matrix}$ ),由于盲人手指敏感区有限,对最下面的两个点不敏感,对8点盲文“摸读”困难,因此实际生活中一般不用8点盲文,而用6点盲文。盲文中6个点分别用数字1-6表示,如左边一排的第1个点称为1点,依次类推。如图1所示,实际使用时,计算机中用黑点表示凸起的点,朗读时读凸起的点,如 $\begin{matrix} \bullet & \bullet \\ \bullet & \bullet \\ \bullet & \bullet \end{matrix}$ 读作“一三五点”。



图1 六点盲文点位表示

由于 Unicode 的盲文不通用,造成目前计算机中还没有统一的盲文字库,因此目前盲文不能在普通计算机中输入与显示。中国盲文出版社研发了阳光软件,并制作了6点盲文字库,使得在普通计算机上能够接入并使用刻印机、点显仪等专用盲用输出设备。

目前国内盲文编辑软件存在以下问题:

(1)盲文字库专用,非专用软件不能使用盲文字库。

出于知识产权考虑,目前国内盲文字库与软件配套使用,盲文只能使用在其软件中,即若脱离了相应软件,盲文无法输入、显示与使用。

(2)没有安装盲文字库的计算机无法显示盲文。

使用点显器、刻印机等设备时,其计算机必须安装盲文字库,这点是毋庸置疑的。但当编辑盲文文档后,尚不能在没有安装盲文字库的计算机显示盲文的点位,这样不便于盲文文档的存储、传输。

(3)盲文与汉字之间转换的正确率无法达到100%。目前我国采用现行盲文,每个汉字用两方盲文表示,一方表示声母,一方表示韵母。如声母 m 用盲文 $\begin{matrix} \bullet & \bullet \\ \bullet & \bullet \end{matrix}$ 表示,韵母 ei 用盲文 $\begin{matrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{matrix}$ 表示,所以汉字“美”拼音为 mei,盲文表示为 $\begin{matrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{matrix}$ ,当然汉字“没”、“每”也都是用同样的盲文表示。现行盲文表示汉字时的固有缺陷使得盲汉或汉盲的转换正确率无法达到100%,效果比较好的也只能达到97%<sup>[1,11-23]</sup>。而盲文阅读是摸读,即盲人依次摸一个个盲文点字,识别成声母、韵母后转为汉字或词组,不能像普通人阅读语句那样能够依据上下文理解,盲人对盲文的容错率低。虽然转换正确率已经令人很满意,但是盲文阅读习惯与汉字不同,就算是个别盲文点字的错误也会导致盲人无法准确理解语义。拥有97%的准确率的盲汉转换系统没有得到实际应用就是因为这个原因。

(4)无法实现盲文与汉字、各种符号的混排与转换。

盲文表示的信息除了文字以外,还能表示数学、物理、化学、音乐等其他学科的符号或者表达式。前面的研究实现了盲文与汉字的转换(翻译)、数学符号与盲文的翻译转换<sup>[24]</sup>,未能实现与其他学科符号的混排、存储、转换。

### 1.3 本文创新点

本文是在我们前期已有研究成果<sup>[16-18]</sup>的基础上,为解决上述问题,开发了盲文文档编辑转码系统,其核心思想是将盲文转换为汉字点位编码表示,从而实现在安装有盲文字库的计算机上显示盲文,能够使用盲文刻印机、点显仪等盲用外设;在没有安装盲文字库计算机上显示为汉字点位编码,便于文档保存、传输;该文档拷贝到第三台装有盲文字库的计算机(其盲文字库可能还不兼容)上能够实现将汉字点位编码再转

换成盲文。

主要创新点:1)最先提出将盲文转为汉字点位编码进行输入、存储、传输,转换后与盲文字库无关;2)使用不确定有穷自动机描述盲文与汉字点位编码的转换过程,并采用逆序拆分子集法对不确定有穷自动机确定化,给出盲文转为汉字点位编码的算法。

## 2 不确定的盲文转码有穷自动机

### 2.1 盲文与汉字转码设计

使用盲文时,可以用盲文的点位表示。例如 $\begin{matrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{matrix}$ 表示拼音“beng(第一声)”,实际使用过程中也可以按照其点位,读作“一二三四五六点一点”。 $\begin{matrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{matrix}$ 在数学领域表示为“1”,读作“三四五六点一点”,本系统就是采用这个特征,实现盲文与汉字码的直接互转。

### 2.2 不确定的盲文转码有穷自动机

定义1 不确定的盲文转码有穷自动机  $M$ (见图2)是一个五元组<sup>[25-29]</sup>: $M=(X, \Psi, f, S, Z)$ ,其中  $X$  是一个有穷状态集, $X=\{X_i, X_i \in C\}$ , $C=\{0, 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16\}$ ;  $\Psi$  是一个输入的有穷汉字表, $\Psi=\{-, 二, 三, 四, 五, 六, 点, 方, 满, 缺, 空\}$ ,为简化表达,用字母依次替代(见图2), $D=\{a, b, c, d, e, f, p, q, m, n, k\}$ ,即  $\Psi=\{\psi_i, \psi_i \in D\}$ ,其中  $D_1=a, D_2=b, D_3=c, D_4=d, D_5=e, D_6=f, D_7=p, D_8=q, D_9=m, D_{10}=n, D_{11}=k$ ;  $f$  为转换函数, $f(X_i, \psi_i) = X_j (X_i, X_j \in X, \psi_i \in \Psi)$ ;  $S$  为初始状态集, $S=\{0\}$ ;  $Z$  是终结状态, $Z=\{6, 7\}$ 。

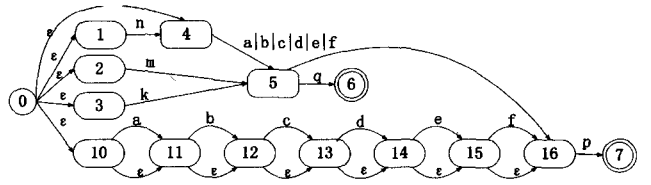


图2 简化不确定的盲文转码有穷自动机(NFA M)

其实  $f$  是一个从  $X \times \Psi^*$  到  $X$  的子集的映射,如语义表达“一三点”和“三一点”可以是等价的, $f: X \times \Psi^* \rightarrow 2^X, 2^X$  表示  $X$  的幂集。但是为了降低转码难度并提高转码的正确性,对64个盲文的汉字转码做出唯一性限制。表1所列为64个盲文与汉字编码的对照表,其中超过5个点(含)用缺方表示,如 $\begin{matrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \end{matrix}$ 表示为“缺六方”,不表示为“一二三四五点”,其目的是减少盲文点位汉字数,使得表示更加简洁。

针对上述的唯一性,对不确定的盲文转码有穷自动机做出如下约束规则。

规则1 最大逆匹配原则。

若存在  $f(f(X_i, \psi_i), \psi_j) = f(X_i, \psi_k) (i \neq j \neq k)$ ,则选择  $f(f(X_i, \psi_i), \psi_j)$  为唯一的状态迁移。

例如:汉字编码“一二三四点”依据最大逆匹配原则转换为盲文 $\begin{matrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{matrix}$ ,而不转换为“一 $\begin{matrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{matrix}$ ”或者其他形式。

规则2 顺序唯一性原则。

$f(f(X_i, \psi_i), \psi_j) \neq f(f(X_i, \psi_j), \psi_i), i \neq j$  且  $\psi_i, \psi_j \in D_n (n=1, 2, 3, 4, 5, 6)$

$\Psi$  中  $\{a, b, c, d, e, f\}$  的输入顺序是唯一的。

例如:汉字点位编码“一四五点”≠“四一五点”≠“五一四点”≠……

盲文“”的唯一编码“一四五点”和“四一五点”根据最大逆匹配原则,转换为“”。

规则 3 终态不等价原则。

$f(x_i, p) \neq f(x_i, q), x_i \in X$ 。

Z 中的终态 {6,7} 不等价。

例如:“四方”≠“四点”。

利用 3 条约束规则,使得不确定的盲文转码有穷自动机简洁,才能完成 NFA 的绘制,图 2 是针对表 1 的描述的不确定有穷自动机。

表 1 盲文与盲文点位汉字对照表

Table with 8 columns: 点位汉字, 盲文, 点位汉字, 盲文, 点位汉字, 盲文, 点位汉字, 盲文. It lists various Braille characters and their corresponding dot patterns for different Chinese characters.

3 不确定盲文转码有穷自动机的确定化

定理 1 设 L 为一个由不确定的有穷自动机接受的集合,则存在一个接受 L 的确定的有穷自动机<sup>[30]</sup>。

试着采用常规子集法对图 2 所示的 NFA M 进行确定化,发现状态转换函数 f 的个数超过 100,处理起来非常麻烦,已经无法通过画图来表示。

分析约束规则后,发现 NFA M 的一些特征:

- (1) 汉字编码的最后一个字是“点”或者“方”;
(2) 按照最后一个字是“点”或者“方”,往前最大逆匹配即可求得盲文点位。

利用这两个特征,采用逆序拆分子集法对 NFA M 确定化。即对于图 2 中的 NFA M,先对其逆序再拆分,使其简化成 2 个子 NFA M,用子集法(逆序拆分子集法,简称 ROSS)分别对子 NFA M 进行确定化。

具体步骤<sup>[30-32]</sup>如下。

步骤 1 图 2 中的 NFA M 的所有箭头方向逆序,输入状态保持不变,可以发现终结状态 6 和终结状态 7 变成初始状态。

步骤 2 分别以图 2 的 NFA M 的状态 6 和终结状态 7 为初始状态,拆分成子 NFA M1 和 NFA M2;

M1 是一个五元组: M1 = (X', Psi', f', S', Z'), 其中:

X' 是一个有穷状态集, X' = {1, 2, 3, 4, 5, 6};

Psi' 是一个输入有穷汉字表, Psi' = {a, b, c, d, e, f, q, m, n, k};

f' 为转换函数, f'(xi, phi) = xi (xi, xi in X', phi in Psi');

S' 为初始状态集, S' = {6};

Z' 是终结状态, Z' = {1, 2, 3}。

M2 是一个五元组: M2 = (X'', Psi'', f'', S'', Z''), 其中:

X'' 是一个有穷状态集, X'' = {7, 10, 11, 12, 13, 14, 15, 16};

Psi'' 是一个输入有穷汉字表, Psi'' = {a, b, c, d, e, f, p};

f'' 为转换函数, f''(xi, phi) = xi (xi, xi in X'', phi in Psi'');

S'' 为初始状态集, S'' = {7};

Z'' 是终结状态, Z'' = {10}。

步骤 3 用子集法分别对 M1 和 M2 确定化并化简。

子集法是不确定有穷自动机确定化的经典算法,此处不赘述,读者可参考文献[30]。

下面对 M2 进行确定化。步骤如下:

(1) 计算 T0 = epsilon-closure(7) = {9, 7}, 并将 T0 加入未标记集 C 中;

(2) 标记 T0, T1 = epsilon-closure(move(T0, p)) = {10, 11, 12, 13, 14, 15, 16}, 将 T1 加入 C, T1 未标记;

(3) 标记 T1, 分别计算 epsilon-closure(move(T1, a)) = {10}, epsilon-closure(move(T1, b)) = {10, 11}, epsilon-closure(move(T1, c)) = {10, 11, 12}, epsilon-closure(move(T1, d)) = {10, 11, 12, 13}, epsilon-closure(move(T1, e)) = {10, 11, 12, 13, 14}, epsilon-closure(move(T1, f)) = {10, 11, 13, 14, 15}, 其均未出现在 C 中,依次将其标记为 T2, T3, T4, T5, T6, T7, 并加入 C;

(4) 标记 T3, epsilon-closure(move(T3, a)) = {10}, 即为 T2, 已经标记;

(5) 标记 T4, epsilon-closure(move(T4, a)) = {10}, epsilon-closure(move(T4, b)) = {10, 11}, 即为 T2 T3, 已标记;

(6) 标记 T5, epsilon-closure(move(T5, a)) = {10}, epsilon-closure(move(T5, b)) = {10, 11}, epsilon-closure(move(T5, c)) = {10, 11, 12}, 即为 T2 T3 T4, 已标记;

(7) 标记 T6, epsilon-closure(move(T6, a)) = {10}, epsilon-closure(move(T6, b)) = {10, 11}, epsilon-closure(move(T6, c)) = {10, 11, 12}, epsilon-closure(move(T6, d)) = {10, 11, 12, 13}, 即为 T2 T3 T4 T5, 已标记;

(8) 标记 T7, epsilon-closure(move(T7, a)) = {10}, epsilon-closure(move(T7, b)) = {10, 11}, epsilon-closure(move(T7, c)) = {10, 11, 12}, epsilon-closure(move(T7, d)) = {10, 11, 12, 13}, epsilon-closure(move(T7, e)) = {10, 11, 12, 13, 14}, 即为 T2 T3 T4 T5 T6, 已标记。

至此,算法终止构建了8个子集( $T_0, T_1, T_2, T_3, T_4, T_5, T_6, T_7$ )。为便于书写, $T_0, T_1, T_2, T_3, T_4, T_5, T_6, T_7$  分别为状态0-7。构造 NFA  $M_2$  的确定的有穷自动机 DFA  $N_2 = (X, \Psi, f, S, Z)$ 。其中  $X = \{0, 1, 2, 3, 4, 5, 6, 7\}, \Psi = \{a, b, c, d, e, f, p\}, S = \{0\}, Z = \{2\}$ ,  $f$  的映射较多,此处不再罗列,具体如图3所示。此处需要说明的是  $f(2, a|\epsilon) = 2$ ,其中  $\epsilon$  是后期调整加上的,因为在拆分图2中的 NFA  $M$  时,损失了  $f(5, \epsilon) = 7$ ,所以为保证完备性,增加了  $\epsilon$ 。

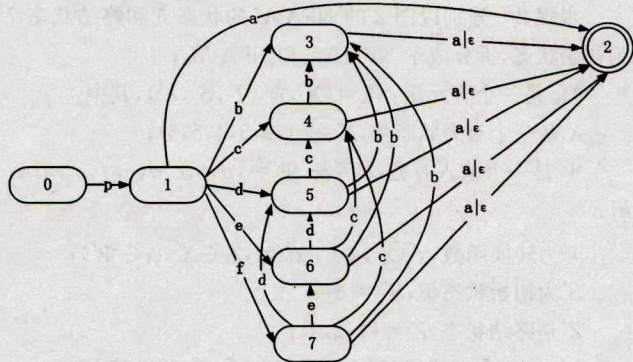


图3 NFA M2 确定化后的 DFA 状态转换图

采用同样步骤对  $M_1$  确定化并化简,得到相应的 DFA  $N_1$ ,如图4所示。

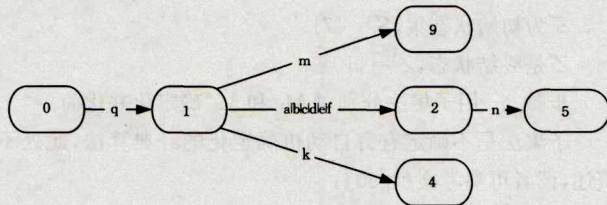


图4 NFA M1 确定化后的 DFA 状态转换图

### 4 实现盲文与汉字编码互转算法

#### 4.1 盲文到汉字点位编码转换

盲文共64个,将每个盲文看成独立的符号,其每个点字具有独立性,与上下文无关,所以可以直接将其替换成对应的汉字编码。

#### 4.2 汉字点位编码到盲文转换

本文重点研究汉字编码到盲文的转换。下面给出汉字编码到盲文转换的算法伪代码描述:

```

(1)建立一个空文本串 S1;
(2)建立文本串 S2,该文本串包含汉字、符号、英文和盲文的汉字编码(如“一三五点”);
for(i=1, i++, i<=S2.Length)
{
    从 S2 中依次取出字符 ai 存入 S1(S1=S1+ai);
    if(ai=="点")
    {
        //按照图3所示的 DFA 状态转换处理
        for(j=5, j--, j<=1)
        {
            从 S1 中取出 a1-j 到 ai 的字符串 S3;
            如果 S3 满足图3所示的 DFA,将 S3 转换为对应的盲文点字的 unicode 的内码;
            否则继续
        }
        if(ai=="方")
    }
}

```

```

//按照图4所示的 DFA 状态转换处理
for(j=3, j--, j<=1)
{
    从 S1 中取出 a1-j 到 ai 的字符串 S3;
    如果 S3 满足图4所示的 DFA,将 S3 转换为对应的盲文点字的 unicode 的内码;
    否则继续
}
其他情况,不处理,继续循环;
}
}

```

### 4.3 实验演示

#### (1)可行性测试

##### 1)中文、盲文混排,盲文与汉字点位编码互转

图5的第一行是中文,下面是对其翻译的盲文表达。图6中的第一行是中文,下面是对其翻译的盲文用点位编码的表达。在系统中选择菜单“汉盲互转”,图5转成图6效果,反之图6转成图5效果。

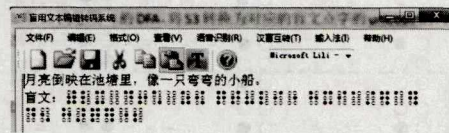


图5 中文与盲文混排

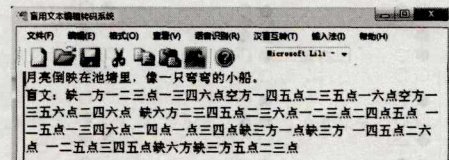


图6 中文与汉字点位编码

##### 2)数学表达式(英文字母)与盲文混排,盲文转为汉字点位编码互转

图7示出了英文、数学符号和盲文混排,并进行盲文与汉字点位编码的互转测试。图7转成图8效果,反之图8转成图7效果。

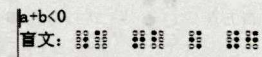


图7 数学式与其盲文表达

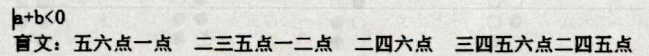


图8 数学式与其盲文汉字点位编码

#### (2)鲁棒性测试

特意设置一些盲文汉字点位编码的干扰(见图9),测试其转换效果。图9中干扰编码有:“三点建议”、“缺一三方”、“五一三四六点”、“四一五点”、“满方满空方”、“缺五点”。图10是图9的转换效果,基本满足转码需求。

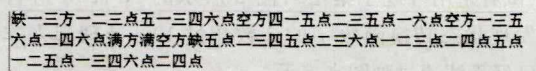


图9 有干扰的盲文汉字点位编码

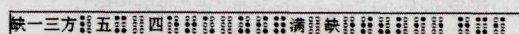


图10 有干扰的盲文汉字点位编码转码结果

上述实验能够证明系统是可行的,即系统能够将盲文与汉字点位编码互转,并能在未安装盲文字库的计算机上使用汉字点位编码存储、传输。

### (3) 准确率测试

输入一段含有汉字、数学符号、音乐符号、盲文的文本串,进行盲文和汉字点位编码的转换,转换正确率接近100%。

对本校5份盲文期末试卷汇总成一个文件,进行转换实验。经统计,5份期末试卷共含37683个字(包括汉字、盲文点字和标点符号),其中包括盲文点字9322个。

#### 1) 盲汉转换实验

转换完成后,经统计,9322个盲文点字全部转换为汉字点位编码,正确率为100%,剩余28361个汉字及符号保持不变。

#### 2) 汉盲转换实验

将前一步已经转换为汉字点位编码的盲文试卷再转换回原始盲文试卷。转换完成后,经统计,34594个汉字(盲文汉字点位编码)被转换为9322个盲文,正确率为100%。

其中,例如句子“我们下午三点在校门口见面”被转化为“我们下午 $\text{\textcircled{3}}$ 在校门口见面”其他如“一点点心意”、“大胆一点的想法”等句子中的“一点”都被转为盲文,共6个,故正确率为 $(28349/28361) \times 100\% = 99.96\%$ 。

为了避免本不属于盲文点位编码的汉字被转换为盲文的问题发生,可以把“三”修改为“3”或者“三点”(中间加一个空格),将其与盲文的汉字点位编码区别开。当然,如果已经发生误转换,也并不影响盲人“听读”内容。

**结束语** 前期已有研究成果<sup>[16-18]</sup>实现输入、存储、输出盲文,已经设计出高效的盲文输入法,也解决了盲文与中文、英文等文字符号的混排,但这些工作必须在安装有盲文字库的计算机中进行,没有安装盲文字库的计算机无法显示盲文,不同的盲文字库也会发生兼容性冲突。

对于本文所提出的将盲文转为汉字点位编码的方法,国内尚无这方面的研究与报道,转换后用汉字点位编码进行输入、存储、传输,理论上与盲文字库无关,可以兼容国内不同盲文字库(对应不同的Unicode码的值);使用非确定有穷自动机描述盲文与汉字点位编码的转换过程,并采用逆序拆分子集法对非确定有穷自动机确定化,给出盲文转为汉字点位编码的算法。

该系统的功能与特点如下:

1) 采用非确定有穷自动机实现盲文与汉字点位编码相互转换,转换正确率接近100%;

2) 盲文能够在安装盲文字库的计算机上使用,转换成汉字点位编码后能在没有安装盲文字库的计算机上存储与使用;汉字点位编码还能够作为不同盲文字库的中间码,解决了不同盲文字库的兼容性冲突问题;

3) 数学、物理、化学、音乐等其他学科的符号也能被转换为汉字点位编码进行保存,盲人可以通过语音(不需要盲文字库支持)听出盲文点字的码位,从而实现盲文文档的阅读,以前的系统在没有安装盲文字库的计算机环境下是无法阅读这类文档的;

4) 在没有安装盲文字库的计算机上录入盲文的汉字点位编码,然后在安装盲文字库的计算机上转换成盲文,并对盲文显示和使用;

5) 实现盲文和普通文字(如中、英、数学符号等)的混排,转换时仅将盲文转换,基本不影响其他文字符号。

将盲文作为标准字库加入现有计算机系统字库中才能彻底解决这类问题。本文所提出的方法对其他非标准字库(自制字库)转码转存也提供了一定启发。

## 参考文献

- [1] WANG Xiang-dong, XIE Shu-hua, ZHANG Bo-ning, et al. Research and implementation of braille input system[J]. Information Technology Letter, 2009(3): 8-16. (in Chinese)  
王向东, 谢书华, 张博宁, 等. 盲文点字输入系统研究与实现[J]. 信息技术快报, 2009(3): 8-16.
- [2] 中国残疾人联合会. 2010年末全国残疾人总数及各类、不同残疾等级人数[EB/OL]. [2012-06-26]. [http://www.cdpf.org.cn/sytj/content/2012-06/26/content\\_30399867.htm](http://www.cdpf.org.cn/sytj/content/2012-06/26/content_30399867.htm).
- [3] ZHANG Ju-xiao. Design And Implementation Of Embedded Input System Of Braille By Using Single-Hand[J]. Computer Technology And Development, 2015(1): 183-186, 190. (in Chinese)  
张居晓. 嵌入式盲文单手输入系统设计与实现[J]. 计算机技术与发展, 2015(1): 183-186, 190.
- [4] LI Nian-feng, DONG Ying-hong, XIAO Zhi-guo. Study on image processing based braille automatic identification system[J]. Manufacturing Automation, 2012, 34(3): 63-67. (in Chinese)  
李念峰, 董迎红, 肖志国. 基于图像处理的盲文自动识别系统研究[J]. 制造业自动化, 2012, 34(3): 63-67.
- [5] ZHANG Shan-jun, YOSHINO K. A Braille Recognition System by the Mobile Phone with Embedded Camera[C]//International Conference on Innovation Computing, Information and Control. Kmamoto, Japan, 2007.
- [6] AL-SALMAN A M S, EL-ZAART A, GOMAI A. A New Approach for Adjusting Braille Image Skewness in Optical Braille Recognition[J]. Communications in Computer & Information Science, 2011, 179: 735-746.
- [7] AL-SALEH A, EL-ZAART A, ALSALMAN A. Dot Detection of Optical Braille Images for Braille Cells Recognition[M]//Lecture Notes in Computer Science. 2008: 821-826.
- [8] JIANG Ming-hu, ZHU Xiao-yan. Segmentation of Mandarin Braille word and Braille translation based on multi-knowledge[C]//Proceedings of ICSP 2000. 2000: 2070-2074.
- [9] JIANG M H, ZHU X Y, Geoges G, et al. Braille to print translations for Chinese[J]. Information and Software Technology, 2002, 44(2): 91-100.
- [10] ZHOU Chun-geng, ZHANG Bing-quan, HUANG He-yan. The Research Andrealization Of Braille-Chinese Machine Translation System Based on Mixed Language Model[J]. Computer Engineering and Applications, 2003, 39(4): 127-130. (in Chinese)  
周春耕, 张秉权, 黄河燕. 基于混合语言模型的盲文机器翻译系统的研究与实现[J]. 计算机工程与应用, 2003, 39(4): 127-130.
- [11] BAO Ta, ZHU Xiao-yan. Research and Implementation of Transformation System Between Mandarin Braille and Chinese[J]. Computer Engineering, 2004, 30(20): 45-46, 100. (in Chinese)  
包塔, 朱小燕. 盲汉转换系统的研究与实现[J]. 计算机工程, 2004, 30(20): 45-46, 100.
- [12] SHAHBAZKIA H R, SILVA T T, RUI M G. Automatic Braille

- Code Translation System[M]//Progress in Pattern Recognition, Image Analysis and Applications, 2005; 233-241.
- [13] FAHIEM M A. A Deterministic Turing Machine for Context Sensitive Translation of Braille Codes to Urdu Text[M]//Combinatorial Image Analysis. Springer Berlin Heidelberg, 2008; 342-351.
- [14] LI Hong-qiao, FAN Xiao-zhong, LI Liang-fu, et al. Feng Research And Implementation of Chinese-Braille Machine Translation System[J]. Computer Applications, 2002, 22(11): 3-6. (in Chinese)  
李宏乔, 樊孝忠, 李良富, 等. 汉语-盲文机器翻译系统的研究与实现[J]. 计算机应用, 2002, 22(11): 3-6.
- [15] YANG Chao, CHE Lei. Research Of transformation System Between Mandarin Braille And Chinese[J]. Journal of Beijing Institute of Graphic Communication, 2011, 19(6): 36-38. (in Chinese)  
杨潮, 车磊. 汉字盲文转换系统的设计[J]. 北京印刷学院学报, 2011, 19(6): 36-38.
- [16] ZHANG Ju-xiao. Study And Implementation of A Braille Plane Inputting System Based on The Private Character Editor[J]. Computer Applications and Software, 2008, 25(4): 284-285. (in Chinese)  
张居晓. 基于造字程序的布莱尔盲文平面输入系统研究与实现[J]. 计算机应用与软件, 2008, 25(4): 284-285.
- [17] ZHANG Ju-xiao. Applying Permutation Algorithm In Compiling Braille Code-Table[J]. Computer Applications and Software, 2010, 27(9): 200-202, 218. (in Chinese)  
张居晓. 排列算法在盲文码表编制中的应用[J]. 计算机应用与软件, 2010, 27(9): 200-202, 218.
- [18] ZHANG Ju-xiao. Implomentation and Improvement of The Method For Inputting Braille Code By Using Single-Hand[J]. Computer Applications and Software, 2014, 31(5): 76-78. (in Chinese)  
张居晓. 单手键入盲文点符编码输入法的实现及改进[J]. 计算机应用与软件, 2014, 31(5): 76-78.
- [19] DING Li-jun. Research And Development of Braille Input Method[D]. Tianjin: Tianjin University Software College, 2014. (in Chinese)  
丁立君. 盲文输入法的研究与开发[D]. 天津: 天津大学, 2014.
- [20] ZHUANG Li, BAO Ta, ZHU Xiao-yan. The Speech and Natural Language Processing Technique Used in a Software System for the Blinds[J]. Journal of Chinese Information Processing, 2004, 18(4): 72-78. (in Chinese)  
庄丽, 包塔, 朱小燕. 盲人用计算机软件系统中的语音和自然语言处理技术[J]. 中文信息学报, 2004, 18(4): 72-78.
- [21] JIANG Xiao-yan, HU Zuo-jin. Design And Implementation of Braille Display Based on Electrode Stimulation Principle[J]. Journal of Computer Applications, 2014, 34(S1): 334-337. (in Chinese)  
蒋小艳, 胡作进. 基于电极刺激原理的盲文点显器的设计与实现[J]. 计算机应用, 2014, 34(S1): 334-337.
- [22] China Braille Publishing House. METEC Braille Display, Braille's Display[J]. China Internet, 2005, (S1): 49. (in Chinese)  
新华-迈泰克点显器: 盲人的“显示器”[J]. 互联网天地, 2005, (S1): 49.
- [23] 中国盲文出版社. 阳光软件使用说明书[EB/OL]. http://www.cbph.org.cn/Article/ArticleShow.asp?ArticleID=58, [2013-03-13] China Braille Publishing House. http://www.cbph.org.cn/Article/ArticleShow.asp?ArticleID=58, 2013-03-13.
- [24] SUN You-yan. The Research on Automatic Conversion for Chinese Mathematical Braille[D]. Lanzhou: Louzhou University, 2013. (in Chinese)  
孙友艳. 数学公式到中国盲文自动转换研究[D]. 兰州: 兰州大学, 2013.
- [25] PENG Dan, GU Wen-xiang. A general Chinese character finite automaton based on moore machine[J]. Computer Engineering and Applications, 2003, 39(36): 68-69. (in Chinese)  
彭丹, 谷文祥. 基于 Moore 机器的一般汉字有穷自动机[J]. 计算机工程与应用, 2003, 39(36): 68-69.
- [26] ZHANG You-hong, GU Wen-xiang, YIN Ming-hao. Optimization Of Chinese Character F Inite Automaton And The Research Relevnant To Its Algorithm[J]. Journal of Northeast Normal University(Natural Science Edition), 2006, 38(2): 22-26. (in Chinese)  
张友红, 谷文祥, 殷明浩. 汉字有穷自动机的优化及相关算法研究[J]. 东北师大学报(自然科学版), 2006, 38(2): 22-26.
- [27] GU Wen-xiang. Chinese Character Finite Automata Theory[J]. Journal of Computer Research And Development, 1995, 32(9): 21-26. (in Chinese)  
谷文祥. 汉字有穷自动机研究[J]. 计算机研究与发展, 1995, 32(9): 21-26.
- [28] CAI Zeng-yu, GU Wen-xiang. Chinese. Character Input Processing Models Based on Two-way Finite Automaton[J]. Journal of Chinese Information Processing, 2005, 19(3): 40-44. (in Chinese)  
蔡增玉, 谷文祥. 汉字双向有穷自动机的研究[J]. 中文信息学报, 2005, 19(3): 40-44.
- [29] CAI Zeng-yu, LIU Shu-ru, ZHANG Jian-wei, et al. Research on Chinese Character Input Processing Models Based on Fuzzy Finite Automaton[J]. Computer Technology And Development, 2008, 18(3): 89-91, 95. (in Chinese)  
蔡增玉, 刘书如, 张建伟, 等. 汉字模糊有穷自动机的研究[J]. 计算机技术与发展, 2008, 18(3): 89-91, 95.
- [30] ZHANG Su-qin, et al. Compiler Principle (second)[M]. Beijing: Tsinghua University press, 2008; 57-58. (in Chinese)  
张素琴, 等. 编译原理(第2版)[M]. 北京: 清华大学出版社, 2008; 57-58.
- [31] ZHENG Li. Study on Snort Rules Modeling, Finite Automaton Transforming And Merging Algorithms [D]. Xi'an: Xi'an University of Science and Technology, 2014. (in Chinese)  
郑丽. Snort 规则建模及有穷自动机的转化与合并算法研究[D]. 西安: 西安电子科技大学, 2014.
- [32] YIN Chao, LI Da-xing. Design of Dfa Based on Improved Bm Algorithm[J]. Microcomputer Information, 2008, 24(7): 215-216, 236. (in Chinese)  
殷超, 李大兴. 基于改进 BM 算法的确定型有穷自动机的设计[J]. 微计算机信息, 2008, 24(7): 215-216, 236.