ISE-TR-80-14

AN INTEGRATED JAPANESE BRAILLE PROCESSING SYSTEM BASED ON KANJI REPRESENTATION

by

Kazuhiko Nakayama
Akifumi Oikawa
Kozo Itano
Yasuhiko Ogawa
Sadao Hasegawa

April 10, 1980

INSTITUTE
OF
INFORMATION SCIENCES AND ELECTRONICS

UNIVERSITY OF TSUKUBA

An Integrated Japanese Braille Processing System Based on Kanji Representation

by

Kazuhiko Nakayama *, Akifumi Oikawa *, Kozo Itano *,
Yasuhiko Ogawa *, Sadao Hasegawa **

* Institute of Information Sciences and Electronics
University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki 305, Japan.

** School for Blind, University of Tsukuba
Mejirodai 3-27-6, Bunkyo-ku, Tokyo 112, Japan.

This work was supported in part by the Science Information Processing Center, University of Tsukuba, and by the grant of research project No.410205: "Basic research on computer input and output of braille towards the expansion of linguistic life circumstances of the blind" as a part of special project research: "Interdisciplinary and comprehensive study of language competence and speech activity" sponsored by the Ministry of Education of Japan (1977-1979).

Abstract

A braille processing system oriented Kanji representation is presented which enables us to establish complete conversion between Japanese ordinary Kanji and braille writings. The system consists of an intelligent braille station backed up with utilities of the large computer. We built several pieces of braille equipment originally for our system such as a braille data typewriter, a braille pin display terminal, and a braille line printer. We implemented an interactive braille text processing system which enables the visually handicapped to have access to ordinary character printing. We also proposed on-line braille dictionaries in order to improve the tools of reading and writing.

1. Introduction

The six-dot braille currently in use was invented by a visually handicapped Frenchman, Louis Braille, in 1825. In this braille, characters are made up of raised dots in a six-dot cell arranged in two vertical columns. In Japan, the standard braille is based on Kana, Japanese phonetic symbol, in which Arabic numerals and the English alphabet can be mingled with Kana. This Kana braille is very convenient for extensive use by the visually handicapped people, because they can read almost all the letters by phonetics.

However, the Kana braille is not omnipotent to convey full meaning of the ordinary Japanese sentences. It is because an ordinary Japanese sentence is a mixture of Kanji and Kana.* In ordinary Japanese writing, appoxiately 10^4 Kanji (Chinese based character or ideographs) are mingled with about 10^2 Kana characters. Each Kanji represents its own set of meanings, and various combinations of Kanji produce Kanji compound words which have specific readings and meanings. In addition, Japanese language is basically different for the language which can be described in Roman letters mainly in that it contains thousands of elements. For example, 12,000 kind of characters are used in a typical general circulating newspaper. Furthermore, there is no clear delimitation between words within a sentence. In Japanese Kana braille, we cannot represent Kanji directly and can only represent phonetic information. This situation causes

---

* Japanese uses traditional Chinese characters called "Kanji" which are ideographic signs, and two types of "Kana" - Japanese phonetic syllabaries consisting of 47 letters.

trouble when we link the two writing worlds: ordinary and braille. The conversion between Kanji and Kana is quite a difficult problem [3], and we can not perform this conversion casually. The complete conversion of Kana to Kanji requires capability to understand natural language. The reverse conversion, Kanji to Kana, requires moreover a large volume Kanji dictionary file.

In order to avoid the difficult problems above, we use a new method of representing Kanji by braille [2,7,8] devised by S.Hasegawa, himself visually handicapped, for the direct representation of Kanji text. If the Kanji braille becomes widespread among the visually handicapped people, the gap between the two writing systems will become small.

Our objective is to establish a integrated braille processing system in order to improve the reading and writing environments of the visually handicapped people. Through the use of computer technology, many braille tools can be supplied easily. Automatic writing systems and on-line braille dictionaries explained later are examples of them.

We have built and improved the experimental braille system which consists of several original input and output equipment systems [1] and software utilities. We have also built an integrated ordinary Kanji processing system [6], and the data base management system IDEAS/77 [5], which can support our braille system.

2. Japanese Extended Kanji Braille

In Kana, one of 10 consonants (null,K,S,T,N,H,M,Y,R,W) is

combined with one of 5 vowels (A,I,U,E,O). This combination produces 50 basic characters, each of which corresponds to a single cell of Kana braille. Kana also includes special sound symbols to indicate voiced consonants, p-sound, double consonants and contracted sounds. For this purpose, Kana braille uses speciall prefixed codes and represents them by double six-dot cells. The same method is used for Arabic numerals and the English alphabet; special shift codes are used. In this case, the shift code is effective until a non-numeric or non-alphabetical character appears. Details for writing Kana braille are specified by the Japan Braille Inscription Method [4].

The Extended Kanji Braille developed by S.Hasegawa includes direct Kanji representation, which depends upon Kana Braille to a great extent. In Kanji braille, two types of Kana, "Hiragana" and "Katakana", are represented as different characters. Hiragana is represented as a single six-dot cell without any prefixed codes, and Katakana is represented with a specific prefixed shift code. In case of numerals, also two type of numerals, Arabic numerals and "Kanji numerals", are introduced. Codes for special symbols such as the period, comma, question mark and exclamation mark are the same as those of Kana Braille. For Kanji representation, 8 types of prefixed code are introduced, and each Kanji character is represented by 3 or 4 six-dot cells shown in Figure 1. Most of Kanji characters have two readings: "on" and "kun". In order to represent a Kanji character, we use two Kana characters which are the first characters in the Kana sequences of the "on" and "kun" reading

respectively. However, in case some Kanji characters have no "on" reading, we use the first two or three Kana characters in the "kun" reading.

It is a distinctive feature of this system that Kanji is interpreted by using both "on" and "kun" readings. This make it much easier to read and write Kanji since their functional factors often accord with their phonetic pronunciation. People who have lost their sight early in life are not familiar with the shape of ordinary characters, but they usually learn to read Kanji in two ways; the word "sea", for example, is read as "kai" and "umi". Therefore, with this system, it is possible for such people to form Kanji with their existing knowledge and also it is easier for them to understand. When a Kanji cannot be constructed with both "on" and "kun" readings, only one reading of the Kanji is symbolized.

3. Integrated Braille Processing System

3.1 Outline of the system Configuration

We have implemented the Integrated Braille processing system which consists of an intelligent braille station backed up with utilities of a large computer. The outline of the configuration of the system is shown in Figure 5. The station is equipped with two types of interactive braille devices: a braille data typewriter and a braille pin display. The station is itself connected to a large computer as a TSS terminal. By this station, we can input and edit the braille text interactively. We also can transform the braille text into ordinary Kanji character, and output to the medium speed Kanji

printer. The English alphabet text can also be handled in this system.

The host computer system, ACOS800/II, supplies several braille utilities and data bases. In the host system, we have installed a high speed braille line printer and an ultra high speed Kanji printer, and realized both conversion, braille to ordinary and ordinary to braille, at both the software and hardware level.

3.2 Braille Input Output Devices

The developement of braille equipment for communication with a computer is quite insufficient. Therefore, we have made several piesces of equipment experimentally.

(1) Braille Line Printer

A high speed braille line printer prints one braille line by dividing it into three rows: upper, middle and lower. The line printer is controlled by a small computer and able to print out 32 braille cells in each line. On the left side of the paper, the ordinary character is printed in 5 x 7 dot image for the operator who cannot read the braille. Pagination is also controlled automatically by the computer and each page is cut from the long roll paper.

(2) Braille Data Typewriter

We have also made a braille data typewriter which consists of a braille keyboard and braille printer mechanism. In the braille keyboard, six-dot braille keys and some other function

keys such as space, backspace, carrige return and break are installed. The typewriter prints one braille character by dividing it into two columns: right and left three-dot sub-cells. Each sub-cell is printed by use of an octagonal braille printing head shown in Figure 2.

(3) Braille Pin Display Terminal

A braille pin display terminal is designed for an interactive device, and is equipped to display a single line of braille the contents of which can be refreshed at a high speed. In this display line, 32 cells of braille characters can be displayed as shown in Figure 3, and each cell has a cursor pin and selection switch. For people who cannot read the braille, the ordinary characters are displayed in 5 x 7 dot image. The terminal also has the braille keys which resemble that of the data typewriter.

3.3 Ordinary Kanji Input Output Devices

At our computer center, an integrated Kanji processing system has been built for the extensive use of Kanji processing in research, education, and house keeping [6], and we can use them for supporting our braille system.

In our system two types of laser printer has been installed. One is a high speed printer which can output about 10000 different characters on each page in high resolution. The speed of the printer is 0.5 pages per second. The other is medium speed printer which can output about 3000 different characters, but in high resolution, too. Both printers use cut

sheets as ordinary copy machines, and can produce high qualty printing. The high speed printer is connected to a small computer and data is transfer by magnetic tapes. The medium speed printer is connected to the intelligent station and can be used on-line.

The Kanji display terminal is connected to a large computer system as a TSS terminal, and can input and edit Kanji texts which consist about 6000 different Kanji characters.

4. Automatic Translation of Kanji Text into Braille

We use JIS (Japan Industrial Standard) code, for the computer internal representation of Kanji texts. In JIS Kanji code, 6439 Kanji characters are defined. Our Kanji braille, however, only covers a subset: about 3000 characters, which are used most frequently in usual reading and writing. This diference causes troubles during conversion of an ordinary Kanji text into braille. It is a difficult problem to seperate a Kanji sequence into the minimum Kanji compound words uniquely. Therefore, if we have a Kanji text based on the JIS code, the present braille system is not able to handle translation sufficiently.

In order to solve this problem, our system uses an interactive conversion algorithm. In our system, most of translation is done automatically as far as the Kanji characters are covered within about 3000 characters of subset. When a Kanji appears beyond this subset, the system requests the operator's indication about the reading of the Kanji sequence. Then, the operator responds with the Kana of the reading, and

the translation is continued. In this case, the information about the reading is stored in the specific data library, and is used in the later translation.

5. Interactive Braille Text Processing

Automatic writing for the visually handicapped people is a rather important tool for the communication between braille and ordinary writings. Such a communication system has been implemented experimentally as a paper tape based system by S.Hasegawa. We have further developed a text processing system based on utilization of the intelligent braille station.

Our braille text processing system consists of an Braille Interactive Editor (BIE) based on the pin display. BIE allows manuscription of the braille text in a workfile. The text can be put into the workfile from files and from the braille keyboard. The workfile can be changed from the braille keyboard, and it can be saved in a file. The text in the workfile is not necessarily divided into lines, although it can be. A new-line is indicated by a new-line character which is displayed by a special braille mark.

Most of the time, BIE displays a portion of the workfile on 32 cell of pin display. On this pin display, a cursor is also displayed by the position-indicate pin over the display line. The cursor marks the place where things usually happen. The place that the cursor marks is always immediately to the left of the cursor, between the character below the cursor and the character to the left of that one. For example, if text is inserted from the keyboard, it always goes immediately to the

left of the cursor. All commands of BIE should be entered by special function keys on the braille keyboard. If we want to move the cursor within line display, a cell selection switch may be pushed. Line-up and down switches are used for moving the cursor by line unit: 32 cells of braille. Further, by the use of the FIND command, we can move the cursor down to the part of the workfile that matches the text specified as the FIND command.

Two types of DELETE commands can be used in BIE. One of them deletes a single cell marked by the cursor, and the other deletes all right cells of the line from the cell marked by the cursor. In this case, if the cursor marks the leftmost cell of the display, all 32 cells on the display are deleted. BIE also has other commands for handling data transfer between files and workfile, and so on.

BIE has simple formatting function and inconsistency check function of the braille text, in order to avoid braille sequences which are not permitted in the braille inscription rule.

6. On-line Braille Dictionaries

Braille printing is very sparse on the paper compared with the ordinary printing. Therefore, condensed printing such as dictionaries can not be made in a practical form, because it quickly becomes a very large volume of paper. As an experimental case, we have installed the content of a dictionary into the data base system, and tried to make an on-line dictionary. An English-Japanese dictonary is used as a test

case. Through the use of the on-line dictionaries, the visually handicapped can use the dictionary, which they have not had access to previously.

In the ordinary world of Japanese, two type of dictionaries exist; One of them is a Kanji based dictionary in which we can look up a Kanji character in order to get the meaning, reading and compound words. The other is a Kana-based dictionary in which we can look up a word by the phonetic reading in order to get the meaning and corresponding Kanji character. The braille world has the same aspect as the ordinary one. A braille Kanji dictionary can be used for looking up braille Kanji characters and compound words, in order to know the reading and meaning. A braille Kana based dictionary can be used to know the meaning and the corresponding Kanji character. These dictionaries improve the reading and writing environment of the visually handicapped, and also teach Kanji braille to them.

Currently, implementation of the above dictionaries is under way, and will be supported by the data base management system of our large computer.

6. Concluding Remarks

Japanese visually handicapped people have no means to communicate with the ordinary writing world directly. By use of our system, we can partially improve the current environment for them, although our braille system is only an experimental model and many improvements would be necessary for better communication. Developement of our system would supply them much writing information of the ordinary world. In contemporary

printing, a phototypesetting technique based on computer technology has been used widely, and many novels have been processed on-line using Kanji printing. Therefore, it would be possible to translate those books into braille automatically. Further, ordinary Kanji processing field has been investgating practical Kanji input technques. Hence, the braille Kanji input method will provide one of the applications.

Acknowledgements

The author would like to express his sincere thanks to Nippon System Co. and Nippon Typewriter Co. for their helpful cooperation.

References:

[1] K.Itano, A.Oikawa and K.Nakayama, Braille Information Processing System on Large Computer System, Proceedings of the 20th annual conference of Information Processing Society of Japan, 1979. (In Japanese)

[2] Proceeding of International Conference on Computerized Braille Production - Today and Tomorrow, London, May 1979.

[3] Y.Sakamoto, Information Processing of Japanese Braille on a Computer (1) - Coversion between Kanji and Braille, Computer Linguistics 12 (2), 1977. (In Japanese)

[4] Braille Committee of Japan, Japan Braille Inscription Method, 1973. (In Japanese)

[5] IDEAS/77 User's Manual, Scientific Information Processing Center, University of Tsukuba. (In Japanese)

[6] K.Itano, H.Horiguchi, A.Oikawa and K.Nakayama, Kanji Information Processing System at University of Tsukuba (1) - Hardware system equipped with a Laser Printer, Proceeding of the 19th annual conference of Information Processing Society of Japan, 1978. (In Japanese)

[7] S.Hasegawa, Printing of Ordinary Characters Containing Kanji from Braille, Proceeding of the 13th annual conference of Special Education Society of Japan, 1975 (In Japanese)

[8] S.Hasegawa, On Braille Information Processing System, Education and Information, No.224, 1976. (In Japanese)

Figure Captions

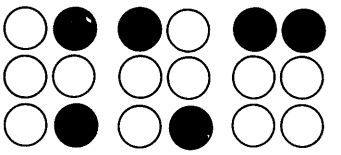
Figure 1. Examples of Braille Kanji Representation by Hasegawa

Figure 2. An Octagonal Printing Head of the Braille Data Typewriter

Figure 3. A Braille Pin Display Terminal

Figure 4. Outline of Braille Processing System

(1) 海 (sea)

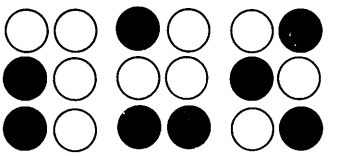


*1 KA U

"on"reading : KAI

"kun"reading: UMI

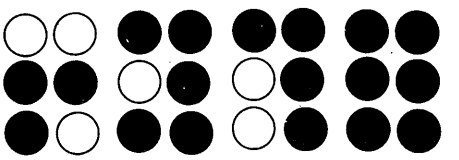
(2) 箱 (box)



*2 HA KO

"kun"reading : HAKO

(3) 娘 (daughter)



*3 MU SU ME

"kun"reading : MUSUME

*1,2,3 are specifying one of the eight Kanji groups.

Figure 1. Examples of braille Kanji Representation by Hasegawa.

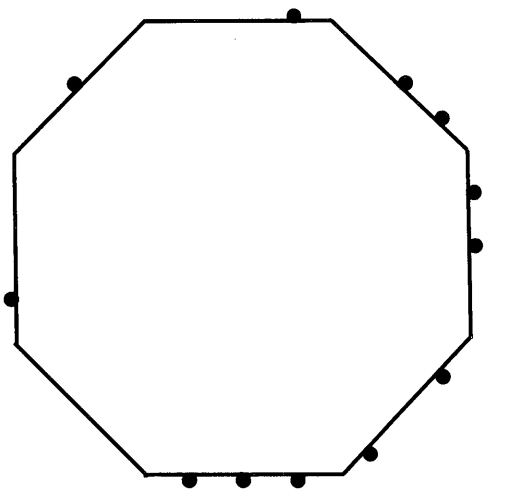


Figure 2. An octagonal printing head of braille data typewriter.

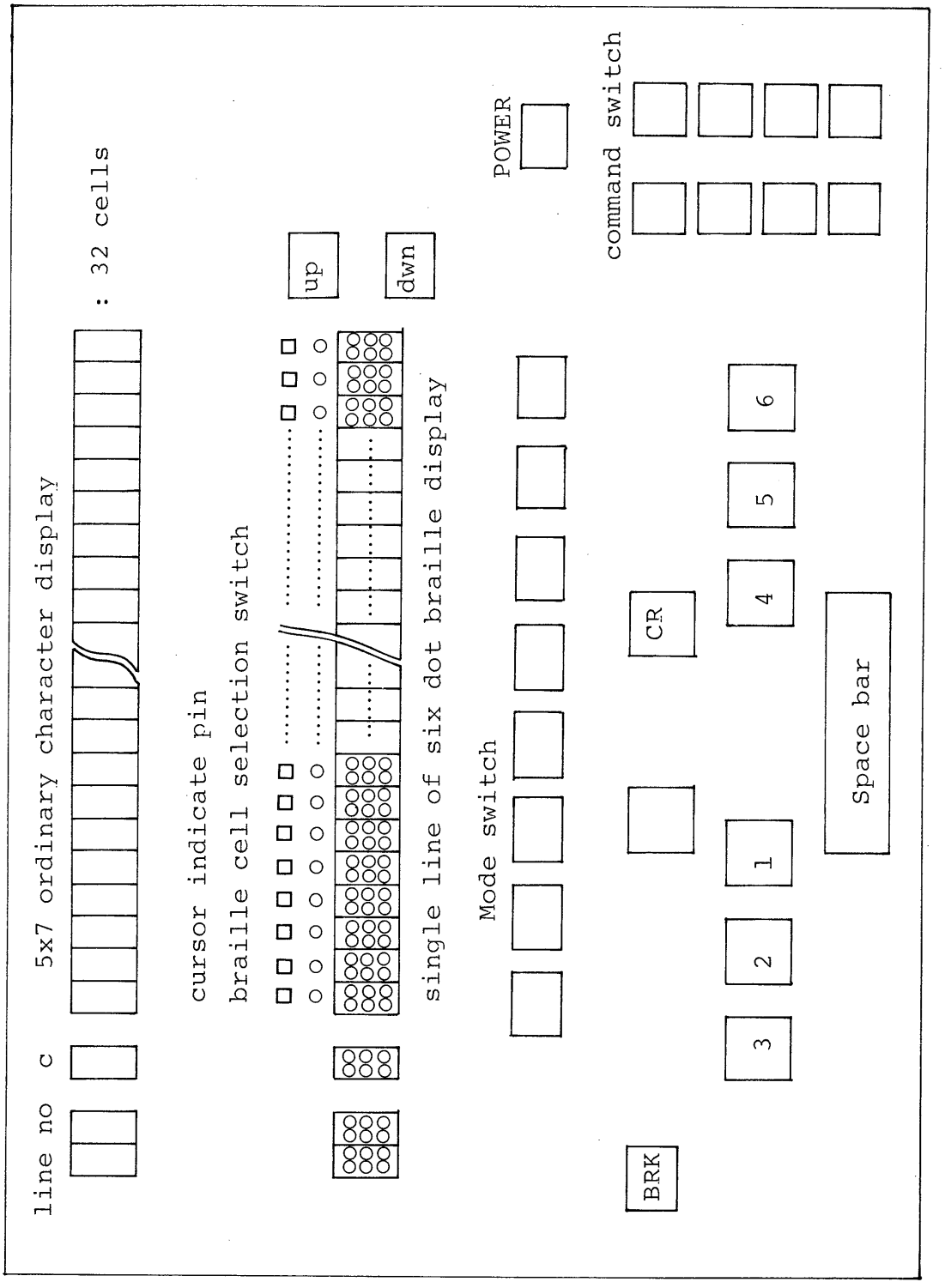


Figure 3. A braille pin display terminal.

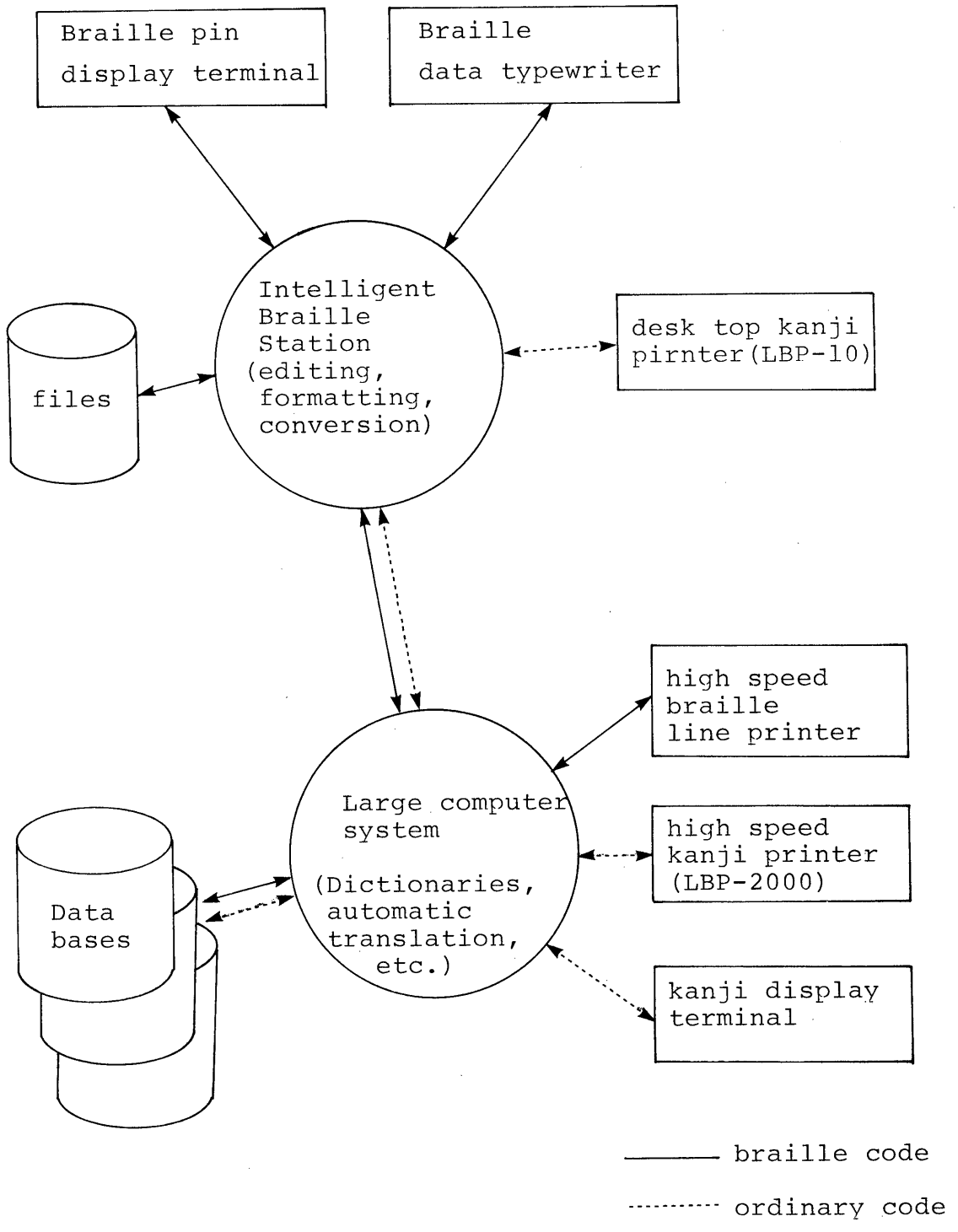


Figure 4. Outline of braille processing system.

INSTITUTE OF INFORMATION SCIENCES AND ELECTRONICS
UNIVERSITY OF TSUKUBA
SAKURA-MURA, NIIHARI-GUN, IBARAKI 305 JAPAN

REPORT DOCUMENTATION PAGE	REPORT NUMBER ISE-TR-80-14

TITLE

An Integrated Japanese Braille Processing System Based on Kanji Representation

AUTHOR(s)

Kazuhiko Nakayama
Akifumi Oikawa
Kozo Itano
Yasuhiko Ogawa
Sadao Hasegawa

REPORT DATE	NUMBER OF PAGES
April 10, 1980	18
MAIN CATEGORY	CR CATEGORIES
management data processing	3.5

KEY WORDS

Blind, braille, text processing, information exchange, braille typewriter, pin display, interactive system.

ABSTRACT

A braille processing system oriented Kanji representation is presented which enables us to establish complete conversion between Japanese ordinary Kanji and braille writings. The system consists of an intelligent braille station backed up with utilities of the large computer. We built several pieces of braille equipment originally for our system such as a braille line printer. We implemented an interactive braille text processing system which enables the visually handicapped to have access to ordinary charcter printing. We also proposed on-line braille dictionaries in order to improve the tools of reading and writing.

SUPPLEMENTARY NOTES