The Development of Unified Japanese Braille Code Based on the Unicode

Mamoru Fujiyoshi[1], Katsuhito Yamaguchi[2], Tooru Ishida[3], Haruhiko Sawazaki[4] and Nobuyuki Ohtake[5], Akio Fujiyoshi[6]

[1] Research Division, National Center for University Entrance Examinations, 2-19-23 Komaba, Meguro-ku, Tokyo 153-8501, Japan, 81-3-5478-1287, fujiyosi@rd.dnc.ac.jp
[2] Nihon University Junior College, 7-24-1 Narashinoidai, Funabashi-shi, Chiba 274-8501, Japan, Eugene@gaea.jcn.nihon-u.ac.jp
[3] National Vocational Rehabilitation Center, 4-2 Namiki, Tokorozawa-shi, Saitama 359-0042, Japan, tooru_ishida@vr.jaed.or.jp
[4] Tokyo Metropolitan Mitaka High School, 6-21-21 Mitaka-shi, Tokyo 181-0004, Japan, haru-s@vm01.vaio.ne.jp
[5] Research Center on Educational Media, Tsukuba College of Technology, 4-12 Kasuga, Tsukuba-shi, Ibaragi 305-0821, Japan, ohtake@k.tsukuba-tech.ac.jp
[6] Department of Computer and Information Sciences, Ibaraki University, 4-12-1 Nakanarusawa, Hitachi, Ibaraki 316-8511, Japan, fujiyoshi@cid.ibaraki.ac.jp

Abstract

It is possible to unify Braille code systems by assigning Braille symbols to "the Unicode," a universal graphic character set for computers adopted all over the world. To the globally used characters in the Unicode such as mathematical symbols, Braille symbols common to all countries are to be assigned. To characters used locally in each country, their own Braille symbols can be defined consistently with the common Braille symbols. Each country can develop their own unified Braille code as the union of the common Braille symbol set and their local Braille symbol set, and those unified Braille code systems are consistent with each other. Actually, based on our methodology, we construct the Unified Japanese Braille Code as the union of "kana," or other local symbols in the present Japanese Braille code, and the common Braille symbols defined in The Unified English Braille Code. The assignment of Braille symbols to primary code blocks in the Unicode is already completed.

Keywords: Braille, visual disabilities, Unicode, Unified English Braille Code, Unified Japanese Braille Code

1. Introduction

As is well known, shortcomings of the current Braille code systems have globally become a serious detriment that hinders people with and without visual disabilities from overcoming barriers between representations in print and in Braille and, therefore, prevents them from sharing a universal information-processing environment (International Council on English Braille, 1995). The current Braille symbols are excessively ambiguous, and the kinds of symbols are definitely insufficient.

The braille code systems were first introduced for the purpose of general writing in local language of

each country. Later, along with the dissemination of education and progress in science and technology, a variety of expansions have been introduced one after the other for the purpose to enable notations in various specialized fields such as arithmetic and mathematics, physics and chemistry, computer programming languages and so on (Fujiyoshi, Ishida, Kizuka, Sawazaki and Yamaguchi, 1996; Fujiyoshi, Ishida, Sawazaki and Ohtake, 2001; Fujiyoshi, Ishida, Sawazaki and Ohtake, 2002; International Council on English Braille, 1995; Ishida, Fujiyoshi, Sawazaki, Yamaguchi and Ohtake, 2002). As a result, even numeral symbols are completely different depending on the country. Furthermore, the Braille code systems in each individual country differ by fields from general texts to those specialized fields. This situation gives rise to serious ambiguity even within local code systems. There is moreover an absolute insufficiency of Braille symbols for notation in the specialized fields.

In the last decade, researches on a unified Braille code system have been in progress in some countries to carry out a drastic reform in Braille code systems. In 1991, the Braille Authority of North America (1992) started a research project called "the Unified English Braille Code (UBC)." Soon after that, many English-speaking countries took interest in it, and "The Unified English Braille Code (UEBC) Research Project" was undertaken as a task by the International Council on English Braille (ICEB) (1995). The final report "Extension of the Base Code" was issued in 1995. The UEBC is expected to be accepted by the English-speaking countries after being adopted by the ICEB in 2004.

In the meanwhile, research on the development of "the Unified Japanese Braille Code (UJBC)" has also been advancing in Japan; its initial version was announced in 1996 (Fujiyoshi, Ishida, Kizuka, Sawazaki and Yamaguchi, 1996). In the UJBC, methodology for the development was clearly defined in a similar way to the UEBC, and Braille symbols were assigned, without any ambiguity, to "non-kanji" code in the Japanese Industrial Standard (JIS) for use in information processing (JIS X-0201 and JIS X-0208) (Japanese Standards Association, 1990).

Incidentally, roughly speaking, Japanese ordinary texts consist of kanji characters (Chinese characters) and non-kanji characters. Main part of the latter consists of kana; they are phonetic symbols to play the same role as alphabet symbols in English.

Here, we propose a systematic methodology for a new approach to the unification of Braille code systems (Fujiyoshi, Ishida, Sawazaki and Ohtake, 2002). It is shown that it is possible to unify Braille symbols internationally by assigning them to the Unicode (The Unicode Consortium, 2000), which is globally adopted as computer text code and, therefore, included as standard code in Windows XP or other software systems. Braille symbols are classified into two categories. Ones in the first category are to be assigned to such characters in the Unicode as are used commonly in printed materials in the world: ones for mathematical expressions, computer-programming languages and so on. Ones in the second category are to be assigned to the characters of Unicode used locally in each language. It is certainly possible to develop a unified Braille code system for each language as a union of the common Braille symbol set and the local Braille symbol set since any inconsistencies do not take place.

Actually, we are now developing the second version of the UJBC based on our new approach (Fujiyoshi, Ishida, Sawazaki and Ohtake, 2001; Fujiyoshi, Ishida, Sawazaki and Ohtake, 2002; Ishida,

Fujiyoshi, Sawazaki, Yamaguchi and Ohtake, 2002). We adopt the Braille symbols of grade 1 mode defined in the UEBC as ones in the first category. They and kana symbols or other local ones in the current Japanese Braille code are all assigned to the graphic characters of Unicode without any conflicts.

This development of the UJBC does substantiate a potential for global unification of Braille code. The UJBC is constructed so as to be completely consistent with the UEBC; that is, those two unified code systems share the same common symbols and also include their own local symbols without any inconsistencies. Furthermore, in the UJBC as well in the UEBC, everything from general texts to technical materials in specialized fields such as mathematics and information processing can thereby be written without any ambiguity. Persons with visual disabilities ranging from elementary school pupils to professional specialists would be able to use the same Braille code system in common at either school or workplace. Furthermore, we strongly believe that the same methodology as in the UJBC could also be applied to contribute to the development of unified Braille code systems for other many languages.

In this paper, we give our methodology to develop the UJBC in the section 2. Next, in the section 3, some samples of specific assignment to the primary code blocks of Unicode are indicated, and character-type attributes are also discussed. The section 4 is conclusion.

2. Methodology for the Development

2.1. Guidelines and rules

The following four guidelines, two subsidiary guidelines and three basic rules are introduced to develop the UJBC based on the Unicode (Fujiyoshi, Ishida, Sawazaki and Ohtake, 2001, 2002; Ishida, Fujiyoshi, Sawazaki, Yamaguchi and Ohtake, 2002).

First, the four guidelines are introduced to show a specific strategy how to assign Braille symbols to the graphic characters of Unicode.

The first guideline is on the fundamental course of assignment; Braille graphic symbols are assigned to the Unicode. The Braille code system generally consists of Braille graphic symbols and Braille indicator symbols; the former is to be assigned to the Unicode. The latter enable to use the same Braille character variously in meaning such as kana, alphabet, numeric and so forth; the assignment to the Unicode is not applied for them since they are original ones in Braille system. If the Braille graphic symbols were assigned to all the necessary graphic characters in the Unicode, it would be possible to express any language in the world by means of the unified system of Braille symbols.

The second is a guideline of sharing same symbols. Independent of the difference in languages, same Braille graphic symbols are to be assigned to Unicode graphic characters if they are used in common throughout the world. The same Braille indicators are also to be shared globally. They form "a common Braille symbol set."

The third is a guideline of originality. To Unicode graphic characters locally used in each language,

their own Braille symbols are to be assigned. In addition, with respect to locally used Braille indicators, their own symbols are also to be defined. They form "a local Braille symbol set."

The fourth is a guideline of union. The total Braille symbol system for each language (each unified Braille code system in other words) is developed as a union of the common Braille symbol set and the local Braille symbol set.

Next, the following two subsidiary guidelines are introduced for the purpose of constructing the UJBC.

The first is the guideline of using 6-dot Braille; in principle, the UJBC uses it as well as the UEBC. If necessary, however, 8-dot Braille could be used. For instance, as far as a refreshable Braille display was concerned, it would be necessary to indicate literal information on a computer monitor.

The second is a guideline of using the kana-based system (Japan Braille Commission, 1990). The local Braille symbols of the UJBC, the same as the current Japanese Braille symbols, are based on the kana system. That is, those symbols are assigned only to kana and other non-kanji symbols in the Japanese sector of Unicode. When it is necessary, however, conversion from kana to kanji can be performed using a function of Japanese word processors.

Finally, concerning a specific assignment and a design of Braille symbols, we introduce the same three basic rules as in the UEBC (Braille Authority of North America, 1992; International Council on English Braille, 1995).

The first is reading rules. A Braille code should be considered primarily from the point of view of the reader, who must understand precisely what symbols are being expressed by a given Braille text. The unified Braille code should be, to the maximum feasible degree, unambiguous.

The second is a transcribing rule. The transcribing should produce Braille that, when the reading rules are applied, yields precisely the original print text (apart from purely ornamental aspects). To make accurate transcription of printed material into Braille, it is absolutely necessary to make sure that the original characters used in the printed material could be precisely specified only from the symbols of the Braille text.

The third is designing rules. The designing of various aspects of a Braille code is an activity that needs to be subject to rules if the code is to maintain direction and coherence over time. That is, the design of the Braille symbol system, including the new definitions or extensions of symbols should be based on definitely established methodology.

2.2. Specific methods for the UJBC development

As is well known, the UEBC consists of the grade 1 mode for notation of technical materials such as mathematical expressions and the grade 2 mode for general writing in which English contractions are used. Since the grade 1 mode is a system that covers almost all necessary internationally common symbols and satisfies our other requirements, we adopt it as the common Braille symbol set. The grade 2 mode of UEBC is regarded as the local Braille symbol set in English. That is, the UEBC is

a union of these two sets.

2.2.1. Introduction of Braille symbol construction rules

There are only 64 Braille characters in 6-dot Braille system. In order to create Braille symbols for all necessary graphic characters in print or other symbols, the UEBC introduces excellent rules to construct Braille symbols as series of Braille characters (Braille Authority of North America, 1992; International Council on English Braille, 1995).

First of all, the Braille characters are classified into 3 groups shown in Fig. 1: a blank cell (a space character), 8 prefixes and the other 55 characters called root characters (or roots, simply).

Figure 1. Three Categories of Braille Characters in the Braille Symbol Construction rules of the UEBC

Space
blank
Eight Prefixes
4, 45, 456, 5, 46, 56, 6, 3456
Fifty-five Root Characters
1, 12, 14, 145, 15, 124, 1245, 125, 24, 245
13, 123, 134, 1345, 135, 1234, 12345, 1235, 234, 2345
136, 1236, 1346, 13456, 1356, 12346, 123456, 12356, 2346, 23456
16, 126, 146, 1456, 156, 1246, 12456, 1256, 246, 2456
2, 23, 25, 256, 26, 235, 2356, 236, 35, 356
34, 346, 345, 36, 3

A UEBC symbol can be defined as a series of zero or more prefixes terminated by a single root (a prefix-root concept). That is, one root character itself, a single prefix plus one root and multiple prefixes plus one root are all regarded as independent Braille symbols. According to this rule, a large number of Braille symbols can be efficiently created without any conflicts. Actually, we can assign Braille symbols to all 1,827 graphic characters in the Unicode, which are primarily needed as the UJBC Braille symbols. Incidentally, we can define 3,912 Braille symbols using series of 3 or less-than-3 cells only.

2.2.2. Union of kana Braille symbols and UEBC symbols

Here, first of all, we have to point out that all kana symbols in the current Japanese Braille Code follow the Braille symbol construction rules in the UEBC (see Appendix A). In addition, with only one exception, serious conflicts do not occur between them and major non-alphabetic symbols in the grade 1 mode of UEBC, such as enclosure symbols, arithmetic symbols and others. Those UEBC symbols can be used directly (without any indicators) in Japanese sentences in Braille.

The one conflict exists; a kana "Gi" (dots 5,126) in Japanese is used as the left parenthesis in the UEBC. It could be solved easily, however, if new alternative symbols for left and right parentheses (dots 6, 126 and dots 6, 345) were defined in the UEBC for the UJBC.

Thus, if only a restricted number of local Braille symbols other than kana were consistently defined, we could obtain the UJBC as the union of UEBC grade 1 mode and local Japanese Braille symbol

set.

2.2.3. Switching between Japanese and International modes

As was mentioned, the grade 1 mode of UEBC is adopted as the common Braille symbol set in the UJBC. Here, we name it "the International mode." In the UEBC, the grade 1 mode is switched from the grade 2 mode by the grade 1 indicators (dots 56 or its series). The grade 2 mode in the UEBC corresponds to the Japanese mode of UJBC in which we can use the Japanese local Braille symbols such as kana. The International mode is switched from the Japanese mode by the International mode indicators that are the same ones as grade 1 indicators. The International mode is completely equivalent to the UEBC grade 1 mode; all symbols are interpreted as ones of UEBC in every sense. As was pointed out previously, major non-alphabetic symbols such as enclosure symbols, arithmetic symbols or others also can be used directly (in other words, without the International mode indicators) in the Japanese mode.

3. Specific assignment of Braille symbols to the Unicode

3.1. Graphic characters subject to assignment

In the second version of UJBC, both of all graphic characters listed in the grade 1 mode of the UEBC and all local ones included in the Japanese sectors of Unicode except for kanji are subject to assignment. In Table 1, we show 33 code blocks of Unicode that are supposed to be assigned. A number of graphic characters contained in them and a number of the characters assigned already for Braille symbols are also given. In total, there are 3,252 characters that are candidates for assignment; however, many of them do not necessarily require Braille definitions in the UJBC. For instance, at counting this number we exclude many Latin characters and some Korean characters included in these code blocks, which are manifestly not used in English and Japanese.

The assignment of 1,827 characters has been completed so far. In particular, the assignment of Braille symbols to all characters in the eight primary code blocks: Basic Latin, Latin-1, Mathematical Operators, Miscellaneous Technical, Enclosed Alphanumeric, Braille Pattern, Hiragana and Katakana (two different classes of kana characters as is described later) has already been done.

3.2. Assignment to the Basic Latin

As samples of our assignment, the Braille symbols and their character-type attributes assigned to the Basic Latin code block are shown in Table 2. It should be noted that the majority of them can be written without any indicators in both of English and Japanese.

"Character type" in headings of character-type attributes at the top of Table 2 indicates three types of Braille character sets based on the following classification. "E" is a character type that has already been defined in the UEBC. "J" is one defined locally in the UJBC such as kana, Japanese punctuation marks and so forth. "X" is a Latin character type; but, it is tentatively defined in the UJBC since it is not currently the UEBC symbol. All of the 95 characters in the Basic Latin have already been defined in the UEBC and, therefore, are of the type E.

"Number of cells" is the number of cells required to represent the Braille symbols in that code block. In the case of Basic Latin, taking account of their high frequency, all of the Braille symbols consist of one or two cells.

"Need for mode indicator symbol" shows an attribute whether or not a mode indicator is required to use that Braille symbol in English or Japanese.

Table 1. Code Blocks of Unicode Assigned to the UJBC

Code Range	Block Name	Number of Code	Number Assigned to Braille Code
0020-007E	Basic Latin	95	95
00A0-00FF	Latin-1 Supplement	96	96
0100-017F	Latin Extended-A	128	128
0180-024F	Latin Extended-B	178	114
0300-036F	Combining Diacritical Marks	82	23
0370-03FF	Greek and Coptic	110	68
1E00-1EFF	Latin Extended Additional	252	0
2000-206F	General Punctuation	82	73
2070-209F	Superscripts and Subscripts	28	28
20A0-20CF	Currency Symbols	16	16
20D0-20FF	Combining Diacritical Marks for Symbols	20	1
2100-214F	Letterlike Symbols	58	58
2150-218F	Number Forms	49	44
2190-21FF	Arrows	100	20
2200-22FF	Mathematical Operators	242	242
2300-23FF	Miscellaneous Technical	54	54
2400-243F	Control Pictures	39	39
2460-24FF	Enclosed Alphanumeric	139	139
25A0-25FF	Geometric Shapes	88	12
2600-26FF	Miscellaneous Symbols	109	6
2700-27BF	Dingbats	160	0
2800-28FF	Braille Patterns	256	256
3000-303F	CJK Symbols and Punctuation	52	24
3040-309F	Hiragana	90	90
30A0-30FF	Katakana	94	94
3200-32FF	Enclosed CJK Letters and Months	144	0
3300-33FF	CJK Compatibility	249	0
FB00-FB4F	Alphabetic Presentation Forms	7	6
FE20-FE2F	Combining Half Marks	4	0
FE30-FE4F	CJK Compatibility Forms	28	0
FE50-FE6F	Small Form Variants	26	0
FF00-FFEF	Halfwidth and Fullwidth Forms	170	101
FFF0-FFFF	Specials	7	0
TOTAL		3252	1827

Table 2. Graphic Character, Braille Symbol and Need of Indicators for the Basic Latin Code Block

Code	Code Name	Graphic Character	Braille Symbol (DOT)	Character Type	Number of Cells	Number of Braille Symbol	Need of Indicator	
							English Mode	Japanese Mode
0020	SPACE			E	1	1	G2	K
0021	EXCLAMATION MARK	!	235	E	1	1	G1	K
0022	QUOTATION MARK	"	6 2356	E	2	1	G2	K
0023	NUMBER SIGN	#	456 1456	E	2	1	G2	K
0024	DOLLAR SIGN	$	4 234	E	2	1	G2	F
0025	PERCENT SIGN	%	46 356	E	2	1	G2	F
0026	AMPERSAND	&	4 12346	E	2	1	G2	K
0027	APOSTROPHE	'	3	E	1	1	G2	F
0028	LEFT PARENTHESIS	(	5 126	E	2	1	G2	F
0029	RIGHT PARENTHESIS	)	5 345	E	2	1	G2	K
002A	ASTERISK	*	5 35	E	2	1	G2	K
002B	PLUS SIGN	+	5 235	E	2	1	G2	K
002C	COMMA	,	2	E	1	1	G1	F
002D	HYPHEN-MINUS	-	36	E	1	1	G2	F
002E	FULL STOP	.	256	E	1	1	G1	F
002F	SOLIDUS	/	456 34	E	2	1	G2	K
0030	DIGIT ZERO	0	3456 245	E	2	1	G2	K
0031	DIGIT ONE	1	3456 1	E	2	1	G2	K
0032	DIGIT TWO	2	3456 12	E	2	1	G2	K
0033	DIGIT THREE	3	3456 14	E	2	1	G2	K
0034	DIGIT FOUR	4	3456 145	E	2	1	G2	K
0035	DIGIT FIVE	5	3456 15	E	2	1	G2	K
0036	DIGIT SIX	6	3456 124	E	2	1	G2	K
0037	DIGIT SEVEN	7	3456 1245	E	2	1	G2	K
0038	DIGIT EIGHT	8	3456 125	E	2	1	G2	K
0039	DIGIT NINE	9	3456 24	E	2	1	G2	K
003A	COLON	:	25	E	1	1	G1	F
003B	SEMICOLON	;	23	E	1	1	G1	F
003C	LESS-THAN SIGN	<	4 126	E	2	1	G2	K
003D	EQUALS SIGN	=	5 2356	E	2	1	G2	K
003E	GREATER-THAN SIGN	>	4 345	E	2	1	G2	K
003F	QUESTION MARK	?	236	E	1	1	G2	K
0040	COMMERCIAL AT	@	4 1	E	2	1	G2	K
0041	LATIN CAPITAL LETTER A	A	6 1	E	2	1	G2	K
0042	LATIN CAPITAL LETTER B	B	6 12	E	2	1	GX	K
0043	LATIN CAPITAL LETTER C	C	6 14	E	2	1	GX	K
0044	LATIN CAPITAL LETTER D	D	6 145	E	2	1	GX	K

0045	LATIN CAPITAL LETTER E	E	6 15	E	2	1	GX	K
0046	LATIN CAPITAL LETTER F	F	6 124	E	2	1	GX	K
0047	LATIN CAPITAL LETTER G	G	6 1245	E	2	1	GX	K
0048	LATIN CAPITAL LETTER H	H	6 125	E	2	1	GX	K
0049	LATIN CAPITAL LETTER I	I	6 24	E	2	1	G2	K
004A	LATIN CAPITAL LETTER J	J	6 245	E	2	1	GX	K
004B	LATIN CAPITAL LETTER K	K	6 13	E	2	1	GX	K
004C	LATIN CAPITAL LETTER L	L	6 123	E	2	1	GX	K
004D	LATIN CAPITAL LETTER M	M	6 134	E	2	1	GX	K
004E	LATIN CAPITAL LETTER N	N	6 1345	E	2	1	GX	K
004F	LATIN CAPITAL LETTER O	O	6 135	E	2	1	G2	K
0050	LATIN CAPITAL LETTER P	P	6 1234	E	2	1	GX	K
0051	LATIN CAPITAL LETTER Q	Q	6 12345	E	2	1	GX	K
0052	LATIN CAPITAL LETTER R	R	6 1235	E	2	1	GX	K
0053	LATIN CAPITAL LETTER S	S	6 234	E	2	1	GX	K
0054	LATIN CAPITAL LETTER T	T	6 2345	E	2	1	GX	K
0055	LATIN CAPITAL LETTER U	U	6 136	E	2	1	GX	F
0056	LATIN CAPITAL LETTER V	V	6 1236	E	2	1	GX	F
0057	LATIN CAPITAL LETTER W	W	6 2456	E	2	1	GX	K
0058	LATIN CAPITAL LETTER X	X	6 1346	E	2	1	GX	K
0059	LATIN CAPITAL LETTER Y	Y	6 13456	E	2	1	GX	K
005A	LATIN CAPITAL LETTER Z	Z	6 1356	E	2	1	GX	K
005B	LEFT SQUARE BRACKET	[	46 126	E	2	1	G2	K
005C	REVERSE SOLIDUS	¥	456 16	E	2	1	G2	K
005D	RIGHT SQUARE BRACKET	]	46 345	E	2	1	G2	K
005E	CIRCUMFLEX ACCENT	^	4 26	E	2	1	G2	K
005F	LOW LINE	_	46 36	E	2	1	G2	K
0060	GRAVE ACCENT	`	46 16	E	2	1	G2	K
0061	LATIN SMALL LETTER A	a	1	E	1	1	G2	F
0062	LATIN SMALL LETTER B	b	12	E	1	1	GX	F
0063	LATIN SMALL LETTER C	c	14	E	1	1	GX	F
0064	LATIN SMALL LETTER D	d	145	E	1	1	GX	F
0065	LATIN SMALL LETTER E	e	15	E	1	1	GX	F
0066	LATIN SMALL LETTER F	f	124	E	1	1	GX	F
0067	LATIN SMALL LETTER G	g	1245	E	1	1	GX	F
0068	LATIN SMALL LETTER H	h	125	E	1	1	GX	F
0069	LATIN SMALL LETTER I	i	24	E	1	1	G2	F
006A	LATIN SMALL LETTER J	j	245	E	1	1	GX	F
006B	LATIN SMALL LETTER K	k	13	E	1	1	GX	F
006C	LATIN SMALL LETTER L	l	123	E	1	1	GX	F
006D	LATIN SMALL LETTER M	m	134	E	1	1	GX	F
006E	LATIN SMALL LETTER N	n	1345	E	1	1	GX	F
006F	LATIN SMALL LETTER O	o	135	E	1	1	G2	F

0070	LATIN SMALL LETTER P	p	1234	E	1	1	GX	F
0071	LATIN SMALL LETTER Q	q	12345	E	1	1	GX	F
0072	LATIN SMALL LETTER R	r	1235	E	1	1	GX	F
0073	LATIN SMALL LETTER S	s	234	E	1	1	GX	F
0074	LATIN SMALL LETTER T	t	2345	E	1	1	GX	F
0075	LATIN SMALL LETTER U	u	136	E	1	1	GX	F
0076	LATIN SMALL LETTER V	v	1236	E	1	1	GX	F
0077	LATIN SMALL LETTER W	w	2456	E	1	1	GX	F
0078	LATIN SMALL LETTER X	x	1346	E	1	1	GX	F
0079	LATIN SMALL LETTER Y	y	13456	E	1	1	GX	F
007A	LATIN SMALL LETTER Z	z	1356	E	1	1	GX	F
007B	LEFT CURLY BRACKET	{	456 126	E	2	1	G2	K
007C	VERTICAL LINE	|	456 1256	E	2	1	G2	K
007D	RIGHT CURLY BRACKET	}	456 345	E	2	1	G2	K
007E	TILDE	~	4 35	E	2	1	G2	K

The "English mode" indicates whether or not the grade 1 indicator is necessary within the UEBC. "G1" means that they can be used in the grade 1 mode without the indicator but must be preceded by it when in the grade 2 mode. "G2" is a character type that can be used in both of the grade 1 and grade 2 mode without the indicator. "GX" is a type to be interpreted as a different Braille symbol in each mode. For instance, "t" simply indicates the letter t in the grade 1 mode. On the other hand, in the grade 2 mode, isolated t is a contraction for the word "that" while ones in letter series express the letter t itself. In the case of Basic Latin, there are 40 G2 symbols, 6 G1 symbols and 49 GX symbols; therefore, only 6 symbols require the grade 1 indicator in general texts of English.

The "Japanese mode" indicates whether or not the International mode indicator must precede those symbols within the Japanese mode. "K" type symbols can be used in the Japanese mode without the indicator. "F" is a character type that requires a preceding "International-mode letter indicator," and "F2" is a character type that needs "International-mode word indicator."

In the case of Basic Latin, there do exist 58 symbols of the type K that include 23 upper-case alphabetic characters. Meanwhile, there are 37 symbols of the type F; however, 29 of them are 26 lower-case and 3 upper-case alphabetic characters. It makes sense that these alphabetic characters require the International mode indicator in Japanese sentences. Only 9 of the 32 non-alphanumeric symbols must be preceded by the International mode indicator in the Japanese mode.

As was mentioned previously, only one conflict exists between kana and the 6 enclosure symbols in the UEBC; namely, the symbol of left parenthesis (dots 5, 126; Unicode 0028 in Table 2) is at present the same as kana "Gi" (dots 5, 126; Unicode 304E in Appendix A).

3.3. Assignment to the eight primary code blocks

We already completed the assignment of Braille symbols to the eight major code blocks; the result is briefly shown in Table 3. The headings at the top of Table 3 are all same as Table 2. For instance,

Number of Cells shows that one Braille symbol consist of one cell to the maximum of nine; incidentally, ones of more-than-four cells are of rather low frequency. Here, we discuss briefly on two code blocks next to the Basic Latin and two kana blocks.

In the Latin-1 block (see Appendix B), 74 of the 95 symbols have already been defined in the UEBC. The other, 22 of them, are left for future job of the UEBC. We tentatively give them Braille definitions in the UJBC.

In the Mathematical Operators block (see Appendix C), 139 of the graphic characters have already been defined in the UEBC. On the other hand, the assignment to the other 103 characters is currently given in the UJBC, and we suggest that they should be added in the UEBC. 135 of them can be used directly in the grade 2 mode of UEBC without the grade 1 indicator, and 103 of them can also be put to use in the Japanese mode of UJBC without International mode indicators.

The current Japanese Braille does not distinguish two different classes of kana: hiragana (see Appendix A) and katakana. The UJBC, in principle, follows that manner. The kana characters are phonetic symbols, and all of the 90 hiragana characters have their counterparts of same sound in the 94 katakana characters. Only 4 of the latter to express sound of foreign origin are uniquely included in katakana. We prepare "The katakana indicator" to differentiate katakana characters from hiragana when it is necessary.

Table 3. Character type and Need of Indicators for Eight Primary Code Blocks of Unicode

Code Range	Block Name	Character Type			Number of Braille Symbol									Need of Indicator					
														English Mode			Japanese Mode		
		E	J	X	1	2	3	4	5	6	7	8	9	G1	G2	GX	K	F	F2
0020-007E	Basic Latin	95	0	0	95	0	0	0	0	0	0	0	0	5	44	46	58	37	0
00A0-00FF	Latin-1	74	0	22	27	59	5	0	4	1	0	0	0	20	76	0	32	2	62
2200-22F1	Mathematical Operators	139	0	103	103	29	54	15	29	9	2	0	1	107	135	0	103	11	128
2300-23FF	Miscellaneous Technical	4	0	50	18	4	5	5	10	4	4	4	0	9	45	0	44	0	10
3040-309F	Hiragana	0	90	0	90	0	0	0	0	0	0	0	0	0	0	0	90	0	0
30A0-30FF	Katakana	0	94	0	94	0	0	0	0	0	0	0	0	0	0	0	94	0	0
2460-24EA	Enclosed Alphanumeric	139	0	0	0	0	9	11	97	22	0	0	0	139	0	0	0	0	139
2800-28FF	Braille Patterns	256	0	0	256	0	0	0	0	0	0	0	0	0	256	0	256	0	0

4. Conclusion

We give the systematic methodology to the unification of Braille code systems by means of assigning Braille graphic symbols to the Unicode (Fujiyoshi, Ishida, Sawazaki and Ohtake, 2001, 2002; Ishida, Fujiyoshi, Sawazaki, Yamaguchi and Ohtake, 2002).

In order to substantiate the potential for unification of Braille code systems, the second version of the Unified Japanese Braille Code (UJBC) is developed by combining both of the grade 1 mode of Unified English Braille Code (UEBC) as common Braille symbols and the current Japanese kana Braille code as local ones to the graphic characters in the Unicode (Fujiyoshi, Ishida, Sawazaki and Ohtake, 2001, 2002; Ishida, Fujiyoshi, Sawazaki, Yamaguchi and Ohtake, 2002). As a result, all major print characters used in either Japanese or English texts - ranging from general writing to various technical materials such as mathematics or other specialized fields - can be expressed using the UJBC only.

Actually, 1,827 characters in the primary code blocks of Unicode are already defined while, in total, 3,252 characters are subject to assignment (Table 1). Furthermore, we classify them according to their character-type attributes. It becomes clear that many symbols assigned to the most basic code blocks can be used without the indicator in the both of English and Japanese texts (Table 3).

It should be remarked that this classification is based on the symbol list cited in the 1995 Final Report (International Council on English Braille, 1995). New definitions or changes after that are not taken into account. We are now working with "UEBC Committee 2 Current Working Symbols List" (International Council on English Braille, 2003) to correct our results.

In the UEBC Committee 2 Current Working Symbols List, a correspondence between the grade 1 symbols and the Unicode is shown. We now strongly expect the ICEB to adopt the same methodology as ours and to use the Unicode more practically. We believe that the UEBC symbols should be defined so as to correspond to the Unicode. Especially, it is important to determine which graphic characters in the Unicode really require Braille definitions in the International mode (that is, the grade 1 mode) and to assign Braille symbols to them officially by the ICEB. They will be supposed to form the common Braille symbol set. On the other hand, all characters that would not be treated in the International mode could be regarded as local ones and might be defined in each local language if necessary. In addition, we would like to request the World Blind Union (WBU) to promote the unification of Braille code systems internationally based on our methodology.

The development of UJBC is a certain step to overcome barriers between representations in print and in Braille. Actually, we are now developing a bi-directional automatic transcription system between texts in the UJBC and in print.

Acknowledgement

We are sincerely grateful to Ms. Kyoko Kittaka, a member of the transcribers' circle for English Braille: "Alto," for her great assistance in accomplishing this work.

References

Braille Authority of North America (1992) Unified English Braille Code Project - Objective II: Extension of the Literary Code.

Fujiyoshi, M., Ishida, T., Kizuka, Y., Sawazaki, H. and Yamaguchi, K. (1996) "Development of the unified Japanese Braille Code - Braille Symbols Assignment to the Code for Information Interchange," *Technical Report of the Institute of Electronics, Information and Communication Engineers (Educational Technology)*, ET96-78 pp. 17-26 (abstract in English).

Fujiyoshi, M., Ishida, T., Sawazaki, H. and Ohtake, N. (2001) "Development of the Second Edition of Unified Japanese Braille Code - Assignment Braille Characters to the Unicode," *Technical Report of the Institute of Electronics, Information and Communication Engineers (Welfare Information Technology)*, WIT2001-27 pp. 1-8 (abstract in English).

Fujiyoshi, M., Ishida, T., Sawazaki, H. and Ohtake, N. (2002) "Worldwide Unification of Braille Code Systems by Means of the Unicode," *Technical Report of the Institute of Electronics, Information and Communication Engineers (Welfare Information Technology)*, WIT2002-1 pp. 1-8 (article in Japanese; English abstract is available).

International Council on English Braille (1992) Unified Braille Code Research Project - Extension of the Base Code.

International Council on English Braille (2003) Committee 2 Current Working Version of UEBC Grade 1, Numeric and Assigned Shape Symbol List as of September 29, 2003, http://www.iceb.org/symbsc2.html.

Ishida, T., Fujiyoshi, M., Sawazaki, H., Yamaguchi, K. and Ohtake, N. (2002) "Assignment of Unified Japanese Braille Code on the Unicode and Development of Braille Registration System," *Technical Report of the Institute of Electronics, Information and Communication Engineers (Welfare Information Technology)*, WIT2002-2, pp. 7-12 (abstract in English).

Japan Braille Commission (1990) Japanese Braille Notation Method (1990 version), Japan Braille Commission (article in Japanese).

Japanese Standards Association (1990) Kanji code for information exchange JIS X-0208, Japanese Standards Association (article in Japanese).

The Unicode Consortium (2000) The Unicode Standard Version 3.0, Addison-Wesley, Massachusetts.