

Unicode に対する統一日本語点字記号の割付と 点字登録・管理システムの開発

石田 透*¹ 藤芳 衛*² 澤崎 陽彦*³ 山口 雄仁*⁴ 大武 信之*⁵

*¹ 国立職業リハビリテーションセンター 〒359-0042 埼玉県所沢市並木 4-2

*² 独立行政法人大学入試センター 〒153-8501 東京都目黒区駒場 2-19-23

*³ 都立三鷹高等学校 〒181-0004 東京都三鷹市新川 6-21-21

*⁴ 日本大学短期大学部 〒274-8501 千葉県船橋市習志野台 7-24-1

*⁵ 筑波技術短期大学 〒305-0821 茨城県つくば市春日 4-12

E-mail: *¹ tooru_ishida@vr.jaed.or.jp, *² fujiyosi@rd.dnc.ac.jp, *³ haru-s@vm01.vaio.ne.jp,
*⁴ eugene@gaea.jcn.nihon-u.ac.jp, *⁵ ohtake@k.tsukuba-tech.ac.jp

あらまし 最近コンピュータの世界共通コード体系として Windows 等に標準搭載されるようになった Unicode に対して世界で共通に使われる図形文字には世界共通の点字記号を割り付け、日本語で使われる図形文字には日本語独自の点字記号を定義し、Unicode に対応した統一日本語点字記号第 2 版の開発を行った。Unicode の記号の数は 4 万種以上と膨大であるため、点字記号の割付を行う時、同じ点字記号を異なる Unicode の図形文字に重複して割り付けるおそれがある。このため、点字記号登録・管理システムを開発し、点字記号の一義性を検証した。その結果、22 件の重複をのぞけば矛盾のないことが検証された。

キーワード 点字, 視覚障害, Unicode, 統一日本語点字記号 UJBC, 統一英語点字記号 UEBC

Assignment of Unified Japanese Braille Code on the Unicode and Development of Braille Registration System

Tooru ISHIDA*¹ Mamoru FUJIYOSHI*² Haruhiko SAWAZAKI*³

Katsuhito YAMAGUCHI*⁴ and Nobuyuki OHTAKE*⁵

*¹National Vocational Rehabilitation Center Namiki 4-2, Tokorozawa-shi, 359-0042 Japan

*²Research Division, National Center for University Entrance Examinations
Komaba 2-19-23, Meguro-ku, Tokyo, 153-8501 Japan

*³Tokyo Metropolitan Mitaka High School Shinkawa 6-21-21, Mitaka-shi, Tokyo, 181-0004 Japan

*⁴Nihon University Junior College Narashinodai 7-24-1, funabashi-shi, Chiba, 274-8501 Japan

*⁵Research Center on Educational Media, Tsukuba College of Technology Kasuga 4-12, Tsukuba-shi, 305-0821 Japan

E-mail: *¹ tooru_ishida@vr.jaed.or.jp, *² fujiyosi@rd.dnc.ac.jp, *³ haru-s@vm01.vaio.ne.jp,
*⁴ eugene@gaea.jcn.nihon-u.ac.jp, *⁵ ohtake@k.tsukuba-tech.ac.jp

Abstract As the Unicode has recently been used in the world as a common graphic character set for Windows, we developed Unified Japanese Braille Code (second edition) to assign worldwide common braille symbols on graphic characters of the Unicode and also define unique braille symbols on those of the Unicode employed only in Japan. We developed a braille registration system and checked uniqueness of these braille symbols on the Unicode by it.

Keyword braille, visual disability, Unicode, Unified Japanese Braille Code, Unified English Braille Code

1. はじめに

点字と墨字の障壁を克服し視覚障害者と健常者との統合的情報処理環境を実現するためには現行の点字記号の不備が重大な障害となっている。現行の日本語点字記号は点字記号の生成原理も定められておらず、あ

まりにも多義的で点字記号が絶対的に不足している⁵⁾
⁷⁾¹⁰⁾。情報科学や数学等、専門分野ごとに記号が異なっている。また、数式表記すらも日本と米国とでは全く異なっている¹⁾。

最近コンピュータの世界共通コード体系として Windows XP や JAVA 言語等に標準で搭載されるように

なった Unicode¹¹⁾ に点字記号を対応させれば世界共通の点字記号体系の開発が可能となる。ただし、Unicode のすべての記号に点字を割り付けるのではあまりにも文字数が多くなる。そこで、世界共通に使われる記号（これを共通ブロックと呼ぶ）を定義し、さらに各国固有に使われる記号（これを各国固有のブロックと呼ぶ）を定義することにより、各国で読みやすい点字体系を構成する。UJBC では、共通ブロックの記号に加え、日本語独特の記号を定義する。

2. Unicode 記号に対する点字記号の割付方法

具体的に Unicode の各記号に UJBC（第 2 版³⁾）でどのような点字記号を割り付けるかの具体的な方法を示す。

2.1 UEBC ですでに定義されている記号の割付

UEBC で使われている記号はそれを採用する。これによって、ASCII 94 文字及び多くの数学記号を割り当てることができる。

2.2 日本語特有の記号についての割付

2.2.1 仮名の清音、濁音、半濁音、促音に関しては日本点字を用いる。カタカナとひらがなはモード指示符によって書き分ける。また、単独の濁点、半濁点に関しては、別の記号を用意し、それらが使われている場合はその記号を用いる。

2.2.2 拗音拗濁音、拗半濁音に関しては文章を書く際、日本表記を用いる。つまり「きゃ」は 2 つのコード、「き」「ゃ」で構成されているが、点字で書く場合は「きゃ」に相当する点字を用いる。

2.2.3 その他日本語のカッコ類は UJBC で独自に定義する。

2.3 UEBC で定義されていないが、世界共通に使われると思われる記号の割付方法

2.3.1 基本文字からの派生

(a) 同一形状あるいは字体の違いを示す記号を作成する。たとえばコード 00B5 の Micro Sign はギリシヤ文字 M と同じ形であるがコードが違う。従ってこれを・・・とする。・で類似あるいは字体が少し違うことを表し、・・で元の文字 M を表す。

(b) リバースあるいはインバースに対応する記号を作成する。たとえば 00A1 の Inverted Exclamation Mark を・・・と書き表す。・・がリバース（インバース）を示す記号であり、・は Exclamation Mark である。

(c) 二重あるいは三重を示す記号を定義する。たとえば 301E の Double Prime Quotation Mark を・・・で表す。・は二重の意味づけを与え、・・はコーテーションのコードである。ただし必ずしも・が二重を表す訳ではない。アクセントなどにも用いられている。また、UEBC で先に定義されている記号の中にも・で始まる記号も存在する。

2.3.2 アクセント記号のついた文字について

(a) ある文字にアクセント記号（アクサングラブ: アクサンテギュなど）がついた文字は原則として、それらの記号がついたことを示す。たとえば Unicode 00C0 の Latin Capital Letter A with Grave を点字では・・・と表す。・・は Grave を、・・が大文字 A を表す。

(b) ただし上の例は Unicode で 0041・0300 の 2 文字でも表すことができる。これを点字では・・・と表し違うコードであることを命じできるようにする。

(c) 単独の Grave は 0060 で UEBC で定義されているので・・を用いるが、単独の Diaeres は・・・と記す。・・は No Bracke Space を表す記号である。

(d) Macron など特殊なアクセント記号は頭文字を・・・で挟み、アクセント文字であることを示す。具体的には・・・で Macron を表す。他にも Breveogonek などがある。

2.3.3 合成文字で表せるものは合成して用いる。

2.3.2 のアクセント文字も広い意味では合成であるが、やや特殊なので別の項とした。この項では、いくつかの単独の文字を組み合わせることで 1 つの文字を示すものを取り上げる。

(a) 否定を表す記号はその文字の後に、否定記号を添えて表す。これは数学記号に多く現れる。たとえば 2204 の there does not exist は・・・と表す。・・は exist の記号であり、・・は否定の記号である。

(b) いくつかの部品に分解される記号は、それらを UEBC の合成規則に従って表記する。合成規則には、横並置、縦並置、重ね書き、及び囲みがある。

カッコで囲まれた英数字たとえば 2474 の Parenthesized Digit One は・・・で表す。・・・は丸カッコであり、・・は数 1 である。・・が横並置を表す記号である。従って開きカッコ、1、閉じカッコが 1 文字と認識される

丸囲みの文字、たとえば 24B6 Circled Latin Capital Letter A は・・・と表す。・・が円を表し、・・は A を表す。・・が次の文字を囲むことを表す。

縦並置は例えば・・の下にハイフンなどに用いる。2265 では・・・を用いる。一方 2267 では UEBC で定義されている・・・を用いる。

重ね書きの例としては 222E の contour integral をあげておくと・・・である。・・が integral・・・が円を表し、これが重なっていることを示す。

2.3.4 あまり使われないであろうと思われる記号は・・と・・で囲む。多くの通過記号はこれに割り当てている。たとえば 20A4 の lira sign は・・・とする。

ただし、頻繁に使われる記号に関してはもっと単純な記号を割り当てることもある。

2.3.5 1記号が複数の文字からなる場合、・・と・・で囲む。これは横並置で表すこともできるが、意味が横並置より明確である場合に用いる。たとえば分数2分の1は.....を用いる。ローマ数字もこの方法で記号を作成する。

2.4 その他のモードについて

2.4.1 フルワイド文字に関してはフルワイド文字指示符を使用する。

これらの方法で現在いくつのコードを割り付けたかを表1に示す。

表1：割付方法と方法別点字記号数

総数	3252
定義済み	1827
リバーズ	31
同一あるいは類似書体	70
アクセントなどがついた文字	
Grave	26
Acute	55
Diaeres/s	31
Circumflex	24
Tilde	12
その他 (Macron など)	116
単独のアクセント文字	23
複数文字との組み合わせ	
否定記号	43
横並置	109
縦並置	62
重ね合わせ	40
囲み記号	89
省略語	91
分数	15
ローマ数字	32
その他のシンボル	35
上付下付添え字	31
Fullwide	100

3. Unicode の各ブロックに対する点字記号の割付状況

3.1 割付対象図形文字

Unicode に対応した統一日本語点字記号 UJBC (第2版)を開発するため Unicode の図形文字に点字記号を割り付ける作業を進めている。割付対象は英語の図形文字と日本語の非漢字の図形文字である。割付対象コード・ブロックと対象図形文字数及び点字割付済み文字数を表2に示す。対象図形文字数は合計3252種である。なお、これらのコード・ブロックの一部含まれているハングル文字は対象外とした。

現在そのうちの1827種に割付を試みた。コード・ブロックのうち、Basic Latin, Latin-1, Mathematical

表2：Unicode の割付対象コード・ブロックとブロック別割付対象図表文字数及び割付済み点字記号数

コード範囲	コード・ブロック名	割付対象文字数	割付済み点字記号数
0020-007E	Basic Latin	95	95
00A0-00FF	Latin-1	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	Letter like 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 Alphanumerics	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

Operators, Miscellaneous Technical, Enclosed Alphanumerics, Braille Pattern, Hiragana, Katakana の主要な8ブロックにはすべて点字記号の割付を完了している。

3.2 Basic Latin に対する割付

点字記号割付の例として Unicode の Basic Latin のコード・ブロックに対して割り付けられた点字記号とその文字種属性を表3に示す。

表3の表頭の文字種属性の見出しの「文字種」は点字記号を定義している3種類の点字文字集合を示す。

「e」はUEBCで定義済みの文字種である。「j」は仮名

表 3: BASIC LATIN コード・ブロックの図形文字と割付点字記号及び文字種の属性

コード	コード名	図形文字	点字記号	文字種	点字記号数	指示符の必要性	
						英語モード	日本語モード
0020	SPACE		・	e	1	g2	k
0021	EXCLAMATION MARK	!	・	e	1	g1	k
0022	QUOTATION MARK	"	・ ・	e	1	g2	k
0023	NUMBER SIGN	#	・ ・	e	1	g2	k
0024	DOLLAR SIGN	\$	・ ・	e	1	g2	f
0025	PERCENT SIGN	%	・ ・	e	1	g2	f
0026	AMPERSAND	&	・ ・	e	1	g2	k
0027	APOSTROPHE	'	・	e	1	g2	f
0028	LEFT PARENTHESIS	(・ ・	e	1	g2	f
0029	RIGHT PARENTHESIS)	・ ・	e	1	g2	k
002A	ASTERISK	*	・ ・	e	1	g2	k
002B	PLUS SIGN	+	・ ・	e	1	g2	k
002C	COMMA	,	・	e	1	g1	f
002D	HYPHEN-MINUS	-	・	e	1	g2	f
002E	FULL STOP	.	・	e	1	g1	f
002F	SOLIDUS	/	・ ・	e	1	g2	k
0030	DIGIT ZERO	0	・ ・	e	1	g2	k
0039	DIGIT NINE	9	・ ・	e	1	g2	k
003A	COLON	:	・	e	1	g1	f
003B	SEMICOLON	;	・	e	1	g1	f
003C	LESS-THAN SIGN	<	・ ・	e	1	g2	k
003D	EQUALS SIGN	=	・ ・	e	1	g2	k
003E	GREATER-THAN SIGN	>	・ ・	e	1	g2	k
003F	QUESTION MARK	?	・	e	1	g2	k
0040	COMMERCIAL AT	@	・ ・	e	1	g2	k
0041	LATIN CAPITAL LETTER A	A	・ ・	e	1	g2	k
005A	LATIN CAPITAL LETTER Z	Z	・ ・	e	1	gx	k
005B	LEFT SQUARE BRACKET	[・ ・	e	1	g2	k
005C	REVERSE SOLIDUS	\	・ ・	e	1	g2	k
005D	RIGHT SQUARE BRACKET]	・ ・	e	1	g2	k
005E	CIRCUMFLEX ACCENT	^	・ ・	e	1	g2	k
005F	LOW LINE	_	・ ・	e	1	g2	k
0060	GRAVE ACCENT	`	・ ・	e	1	g2	k
0061	LATIN SMALL LETTER A	a	・	e	1	gx	f
007A	LATIN SMALL LETTER Z	z	・	e	1	gx	f
007B	LEFT CURLY BRACKET	{	・ ・	e	1	g2	k
007C	VERTICAL LINE		・ ・	e	1	g2	k
007D	RIGHT CURLY BRACKET	}	・ ・	e	1	g2	k
007E	TILDE	~	・ ・	e	1	g2	k

や句読点及びカギ類等, UJBC で独自に定義した文字種である。「x」は英語の文字種であり, 現在は UJBC で定義しているけれども将来 UEBC で定義することを推奨する文字種である。Basic Latin の場合, 95 文字ともすべて UEBC で定義済みである。

「記号数」はそのコードの点字記号を定義するために必要な点字記号の数である。Basic Latin の場合は

使用頻度が高いため点字記号はすべて点字 1 マスないし 2 マスの単一の点字記号で定義されている。しかし, 打ち消しの斜線を含む図形文字等, 複数の図形文字の合成で構成されている文字は複数の点字記号によって表記される。

「モード指示符の必要性」はその点字記号を英語モードまたは日本語モードで表記する場合にモード指示

符の必要性の有無を示す属性である。

「英語モード」は英語モード中でグレード 1・モード指示符の必要性の有無を表す。「g1」はグレード 1・モードではそのまま使用可能であるけれどもグレード 2・モード中ではグレード 1・モード指示符を前置する必要がある文字種である。「g2」はグレード 1・モード指示符なしでグレード 2・モードでもグレード 1・モードでもそのまま表記可能な文字種である。「gx」はグレード 1 とグレード 2・モードとは異なる点字記号を表す文字種である。たとえば、t はグレード 1・モード中ではアルファベットの t を表す。一方、グレード 2・モード中では that の略字を表す t を表す。

Basic Latin の場合、「g2」の記号が 40 個あり、グレード 1 でもグレード 2 でも指示符なしでそのまま使用できる。グレード 2・モードでグレード 1・モード指示符を必要とする「g1」の記号はわずか 6 個である。

「gx」の記号は 49 個あり、グレード 1 ではアルファベットを表し、グレード 2 では略字を表す。

「日本語モード」は日本語モード中でその点字記号の表記に英数モード指示符の前置の必要性の有無を示す。「k」は指示符なしで仮名書きモードでそのまま表記可能な文字種である。「f」は英数モード記号指示符を前置する必要がある文字種である。「f2」は英数モード記号ワード指示符を前置する必要がある文字種である。Basic Latin の場合、日本語モード中では英数モード指示符なしで仮名文字と混在してそのまましよう可能な「k」の文字種が 57 個もある。英数モード記号指示符の前置を必要とする「f」の文字種が 38 個ある。しかし、そのうちの 26 個は小文字アルファベットである。また、23 個は大文字のアルファベットである。このため 32 個の特殊記号のうち英数記号モード指示符の前置を必要とする文字種はわずか 8 個にすぎない。また、小カッコ、中カッコ、大カッコの 6 個のカッコ類のうち仮名とかち合うのは仮名の濁音の「ぎ」とかち合う小カッコの開き記号の「・」の 1 つだけである。

Basic Latin のコード・ブロックの多くの点字記号は英語と日本語の文書中で指示符なしでもそのまま表記可能である。

4. 統一点字記号登録・管理システム

4.1 システムの概要

Unicode の記号の数は 4 万種以上と膨大であるため、点字記号の割付を行う時、同じ点字記号を異なる Unicode の文字に重複して割り付けるおそれがある。このため、点字記号登録・管理システムを開発し、点字記号の一義性を検証する。具体的には簡単なチェックを行うことにより、今までの点字記号とのかち合いをチェックすることとした。

システムでは次のようなことを行っている。

(1) コードの並べ替え

現在 UJBC の表は次のような順序で構成している。

Unicode 番号→文字の説明→割り付ける点字記号→フラグ

これを次のように並べ替える

割り付ける点字記号→フラグ→Unicode 番号→記号の説明

(2) 点字記号順に並べ替える

ここで点字記号に関してソート（整列）を行う。すると、もし同じ記号があれば、それらは並んで出現する。

(3) 同じ記号の出力

最後に共通する点字記号をすべて出力する。

4.2 点字登録・管理システムの検証の結果

登録・管理システムを作成する前に、作成していた表に対して、検証を行ったところ、188 件が出力された。ただし、このうち、多くは、英語モードの記号と日本語モードの記号がかち合っていたためである。これらは矛盾していない。（管理システムでこのチェックを組み入れることは可能であるが、フラグの検証をも併せて行うため、あえて出力させてみた）。

これらを取り除いて調べた結果、64 件の記号の衝突がみられた。

これらは現在修正を行っており、現行執筆時点で 22 件が残っている。これらは従来日本語独特の記号と考えていたものが、共通ブロックで使われる可能性があることがわかったが、それらのコードを今現在そのままにしているためである。現在記号を吟味し、かち合いを無くすことができる予定である。

人間が新しい記号を作り出す時、どれだけ配慮しても、記号のかち合いを起こしてしまうことがわかった。今後登録・管理システムができているため、このようなかち合いは起こらない。

また、かち合いが報告された時点でコード設計の方針を確認する手段ともなる。

5. 考察

世界共通コード体系 Unicode¹⁾の共通ブロックと非漢字の日本語の図形文字に対して点字記号を割り付け、統一日本語点字記号 UJBC の第 2 版の設計を進めている。定義対象図形文字の総数は 3252 種である。そのうち、すでに 1827 種の点字記号の定義を完了した。その結果、英語と日本語の文書を作成するために必要とする主要なコード・ブロックの図形文字に対して点字記号の割付を終了した。

今回開発した点字記号登録・管理システムにより点字記号を登録し、UJBC の点字記号の重複等、点字記号の構成に矛盾のないことを検証した。その結果 22 件の

重複を除けば矛盾のないことが検証された。この 22 件の重複は日本語独自の文字種に対する点字記号を変更することによって解消することが可能である。

今後、未定義の 1425 種の図形文字に点字記号を割り付け、UJBC 第 2 版の完成を目指す予定である。また、UJBC 第 1 版の点字と墨字の相互自動変換システムを UJBC 第 2 版用に改良し、実際に電子ファイルを変換して相互変換時の問題点を洗い出し、UJBC の改良に努める。

文 献

- [1] 藤芳衛, “統一日本語点字記号の開発,” 視覚障害, No.131. pp.39-46, 1994.
- [2] 藤芳衛, 石田透, 木塚泰弘, 澤崎陽彦, 山口雄仁, “新しい統一日本語点字記号の開発—情報交換用符号に対する点字記号の割付—,” 電子情報通信学会技術報告 (教育工学), ET96-78. pp.17-26, 1996.
- [3] 藤芳衛, 石田透, 澤崎陽彦, 大武信之, “統一日本語点字記号第 2 版の開発—Unicode に対する点字記号の割付—,” 電子情報通信学会技術報告 (福祉情報工学), WIT2001-27. pp 1-8, 2000.
- [4] 藤芳衛, 山口雄仁, 石田透, 澤崎陽彦, “新しい統一日本語点字記号の開発—(1) 開発理念と具体的設計—,” 日本特殊教育学会第 34 回大会発表論文集, pp.32-33, 1996.
- [5] International Council on English Braille, Unified braille code research project: Extension of the base code, 1995
- [6] 石田透, 藤芳衛, 澤崎陽彦, 山口雄仁, “新しい統一日本語点字の開発—(3) 3 種の点字数学表記の比較—,” 日本特殊教育学会第 34 回大会発表論文集, pp.36-37, 1996
- [7] 日本企画協会, 情報交換用漢文字符号 JIS X-0208, 日本企画協会, 1990
- [8] 日本点字委員会, 日本点字表記法 (1990 年版), 日本点字委員会, 1990
- [9] 澤崎陽彦, 藤芳衛, 石田透, 山口雄仁, “新しい統一日本語点字記号の開発—(4) 点字記号登録システム—,” 日本特殊教育学会第 34 回大会発表論文集, pp.38-39, 1996
- [10] 山口雄仁, 藤芳衛, 石田透, 澤崎陽彦 “新しい統一日本語点字記号の開発—(2) UJBC 導入の一般文書用点字への影響—,” 日本特殊教育学会第 34 回大会発表論文集, pp.34-35, 1996
- [11] The Unicode Consortium, The Unicode Standard Version 3.0, Addison Wesley, Massachusetts, 2000