

[デジタルタイプー文字情報処理基盤の今とこれからー]

5 情報処理基盤としてのフォントの整備 —源ノ角ゴシック・源ノ明朝の開発—



Ken Lunde | CJKV Type Development, Adobe Inc.

開発の経緯と背景

Pan-CJK 書体は、日本で使用される日本語フォントなどのように特定の地域や言語で使用されることを意図した書体とは異なり、中国（中華人民共和国）、台湾、香港特別行政区、日本、韓国（大韓民国）など複数の東アジア地域および言語をすべて対象としている。Pan-CJK 書体の設計と開発の難しさは、各地域の特性や慣例に適應するために、漢字のみならず句読点などもコードポイントごとに複数のグリフが必要になる可能性がある点に起因する。サポートされている言語間で同一グリフを共有できるかどうかは、書体分類上の様式のみならず実際の個別書体デザインによっても

異なる。図-1 は、Pan-CJK 書体において、5つの異なるグリフを含む68の漢字を示している。

Pan-CJK フォントには、それを使用する上で主に2つの利点があると考えられる。1つ目の利点は書体デザインに関係している。具体的には、各種ウエイトをカバーした上で言語間で一貫した書体デザインを示すことができるということ。これはブランディング上のさまざまな目的、特に複数の地域で顧客にサービスを提供する企業にとっては役に立つ。2つ目の利点は、スペースの節約である。Pan-CJK フォントは、同等の地域または言語固有の複数のフォントを用いた場合よりも全体としてはサイズが小さく、グリフ総数も少ないためスペースが節約できる。たとえば源ノ角ゴシック (Source Han Sans) という Pan-CJK フォントは、65,000 のグリフを含みサイズは約 17 メガバイトだが、対応する5つの言語固有のサブセットフォントを合わせたものは、約 116,000 のグリフ数でサイズは約 30 メガバイトであり、グリフの数とファイルサイズの両方で 45% 近くが重複していることになる。

開発の開始時期は筆者が Adobe に入社したばかりのころ、四半世紀前に遡る。筆者は、1994年9月の UIW6 (Unicode/ISO 10646 実装者ワークショップ 6) において、Pan-CJK フォント開発についての最初の発表、「Creating Fonts for the Unicode Kanji Set: Problems & Solutions」という題名のプレゼンテーションを行った。またそのプレゼンテーションにおいて、複数の地域固有の CMap リソースを使用した CID キー付きフォントの実装をデモンストレーションした。元同僚の Dirk Meyer と筆者は、1999年の初めに最初の実用的なプロトタイプを作成し、1999年9月に IUC15 (第15回国際ユニコード会議) で「Unihan Disam-

Simplified Chinese	傑僭割劇區叟喝塌姿羸懣度扇扉搨摩榻 洩潛瀛瘦瞎磨窘竇箭篠筵糙網蠹羸翁翦 翮育羸艘花裯褐謁譜豁羸轄返迷途造週 遍遭選遼鄰覺閼雕靠靡颯飯驎鬣魔麗麟
Traditional Chinese—Taiwan	傑僭割劇區叟喝塌姿羸懣度扇扉搨摩榻 洩潛瀛瘦瞎磨窘竇箭篠筵糙網蠹羸翁翦 翮育羸艘花裯褐謁譜豁羸轄返迷途造週 遍遭選遼鄰覺閼雕靠靡颯飯驎鬣魔麗麟
Traditional Chinese—Hong Kong	傑僭割劇區叟喝塌姿羸懣度扇扉搨摩榻 洩潛瀛瘦瞎磨窘竇箭篠筵糙網蠹羸翁翦 翮育羸艘花裯褐謁譜豁羸轄返迷途造週 遍遭選遼鄰覺閼雕靠靡颯飯驎鬣魔麗麟
Japanese	傑僭割劇區叟喝塌姿羸懣度扇扉搨摩榻 洩潛瀛瘦瞎磨窘竇箭篠筵糙網蠹羸翁翦 翮育羸艘花裯褐謁譜豁羸轄返迷途造週 遍遭選遼鄰覺閼雕靠靡颯飯驎鬣魔麗麟
Korean	傑僭割劇區叟喝塌姿羸懣度扇扉搨摩榻 洩潛瀛瘦瞎磨窘竇箭篠筵糙網蠹羸翁翦 翮育羸艘花裯褐謁譜豁羸轄返迷途造週 遍遭選遼鄰覺閼雕靠靡颯飯驎鬣魔麗麟

■ 図-1 5つの地域における文字の違い

biguation Through Font Technology」という題名のプレゼンテーションを行った。その後、2009年10月に開催されたIUC33(第33回国際化とUnicodeの会議)で、「Designing & Developing Pan-CJK Fonts for Today」という題名で、Pan-CJKフォントの開発について発表した。

このプレゼンテーションがGoogleの興味を引き、2010年にはジョイントプロジェクトの話が持ち上がった。2012年に両社のビジネス要件が合意に達し、2つのPan-CJK書体ファミリーの開発が開始された。Sans Serifファミリーのバージョン1.000は、Adobeブランドの源ノ角ゴシックおよびGoogleブランドのNoto Sans CJKとして2014年7月にリリースされた。Serifファミリーのバージョン1.000は、源ノ明朝(Source Han Serif)およびNoto Serif CJKとして、2017年4月に同様のブランドでリリースされた。その多くの機能強化の中で繁体字中国語の2番目の形式として香港特別行政区のサポートを追加したSans Serifファミリーのバージョン2.000が2018年11月にリリースされた。Serifファミリーのバージョン2.000のための作業が現在進行中である。

AdobeとGoogleは、ビジネスニーズが異なるため、これらのPan-CJK書体を異なる方法で扱っている。AdobeはSource Hanファミリーを完全に独立したスタンドアロンフォントとして扱うが、GoogleはNoto CJKファミリーをNoto(「No Tofu」の省略形)として明示される非常に大きなパズルの中の大きな部分集合として扱っている。「No Tofu」の「Tofu」とは、選択されたフォントファミリーに表示すべき文字が含まれない場合、彼らのエコシステムによりフォントフォールバックが実行され表示される、いわゆるトーフ(空の長方形グリフ)である。AdobeとGoogleは、ビジネスニーズや戦略が異なる大企業であるため、実態は同一のフォントではあるが、ブランディング上の種々の目的のために2つの異なる名前が付けられているのである。

実際の文字デザインに関しては、Adobeのチーフタイプデザイナーである西塚涼子がデザイン全体の監修を担当しているが、Adobeは中国および韓国の書体

デザインの専門知識を持っていないため、これらの地域の主要なフォント開発会社と提携した。中国語(漢字とbopomofo)はChangzhou SinoType(中華人民共和国)、韓国語(ハングルの音節、文字、記号)はSandoll(大韓民国)、韓国語(漢字)と追加の日本語(漢字)にはイワタ(日本)、などである。また、現在中国語に関しては、Arphic Technology(台湾)と協力している。文字デザイン以降のプロセスはすべて、フォント開発において実績と経験のあるAdobe社内で完璧に行われている。これは、Pan-CJKフォントが100万近くのグリフを扱わねばならないという事実からくる必然であると思われる。

これらのPan-CJK書体ファミリーは、Adobeの小塚明朝や小塚ゴシックなど、また同様にIPA((独)情報処理推進機構)フォントなどの従来の日本語書体ファミリーとは異なる位置付けのものである。ここでサポートされている文字は、対象地域である日本、韓国、中国、台湾、および香港特別行政区で最も一般的または必須と考えられている文字である。フォント構造内で各グリフは、そのUnicodeコードポイントまたはシーケンスに従って命名されるため、Unicodeもプロジェクトで重要な役割を果たしている。日本語のカバレッジはややユニークである。Adobe-Japan1-7のグリフセットには、Pan-CJKの利用状況では有用でないと考えられるシンボル等が含まれているため、完全にはサポートされていない。ただし、Adobe-Japan1 IVS(Ideographic Variation Sequences)およびCJK互換漢字に対応するAdobe-Japan1 SVS(Standardized Variation Sequences)を含む、すべての漢字はサポートされている。さらに付け加えるならば、一部のJIS X 0212およびJIS X 0213の文字と記号は意図的に除外されている。

開発の詳細と課題

Source Han書体の持つ大規模な範囲とその特質は、新しい可能性と機能性を探求する機会を提供する

と同時に、OS、アプリケーション、レイアウトエンジン、およびライブラリの側が誤った憶測をしている場合があることを明らかにするという副作用ももたらした。それら問題の一因は、Pan-CJK フォント本来の性質に起因するものだが、同時にこれがオープンソースであることでより多彩で柔軟性のある実験や研究が可能となったことは喜ばしい。実際のフォントの基本的な特徴は、以下のとおりである。

- 1) フォントリソースに最大のグリフ数 65,535 (CID0 から 65534) を含む
- 2) 特別な目的の Adobe-Identity-0 ROS をベースとした CID キー方式の OpenType/CFF フォントとして実装されている (ROS とは、Registry, Ordering, Supplement の略。CIDFont や CMap リソース中に必須の CIDSystemInfo 中の 3 つの値であり、そのうち最初の 2 つ Registry, Ordering は互換性の目的で使用される)
- 3) 開発版やメジャーバージョン間での CID の変更による影響を避けるため、Unicode ベースの作業用グリフ名を使用し、さらに Unicode マッピングおよび OpenType 機能の確立に貢献している
- 4) Pan-CJK であり、したがって、複数の東アジアの言語およびそれらの地域の慣例をサポートする
- 5) クロスプラットフォームの縦書き用メトリクス設定を使用する
- 6) 2- および 3-em (全角) の文字のグリフとその縦書き形式を含んでいるため、フォントのバウンディングボックスの値が異常に大きくなり、行の高さや行間の点で一部のアプリケーションの不具合を誘発する可能性がある
- 7) オープンソース、つまりいかなる種類のライセンス料も無料である

今回の試みにより、フォントをオープンソースとして利用可能にするというパラダイムは、フォントに対する人々の期待に影響を与えることが分かった。それは、同品質の市販のフォントには不具合を指摘したり改良を要求する方法がほぼないか、あるいは開発者側が不

具合の修正や改良を積極的には提供しないことに由来するのかもしれない。

Source Han 書体は「loc」(Localized Forms) GSUB 機能を最初に実装したフォントであり、この機能は句読点と数字に加えて主に漢字を対象として働くものである。この機能の使用には課題があるが、それは今日のアプリケーションでは、文字、段落、または文書レベルでのテキストの言語タグ付けが一般的にサポートされていないことである。現状では、最新のブラウザと Adobe InDesign だけが、適切なレベルの言語タグ付けサポートを提供している。ほかの Adobe のアプリケーションや Microsoft Word もそのサポートを追加している最中である。この現状に対応するために、本来の Source Han フォントには完全に機能する「loc」GSUB 機能を含めつつ、それぞれデフォルトの言語を持つ別々のフォントをリリースすることとした。Source Han Sans バージョン 2.000 は、日本語、韓国語、簡体字中国語、台湾語の繁体字中国語、香港特別行政区の繁体字中国語の 5 つのデフォルト言語でリリースされた。Source Han Serif バージョン 1.000 は、香港特別行政区の繁体字中国語を除いた 4 つのデフォルト言語でリリースされたが、バージョン 2.000 では 5 番目のデフォルト言語として追加される予定である。つまり、現状で各サポート対象言語のデフォルトグリフにアクセスする方法は 2 つある。1) 対象の言語に対応するフォントを選択する。または、2) 文字、段落、または文書レベルでテキストを言語タグ付けする (後者の方法では、言語タグを参照し「loc」GSUB 機能をサポートするアプリケーションが必要である)。将来恐らくバージョン 3.000 をリリースする時点で、6 番目の言語としてマカオ特別行政区の繁体字中国語を追加する予定である。

各ファミリー用の言語固有のフォントは同じグリフセットを共有するため、ウエイト別およびファミリー別の OpenType/CFF コレクション (別名 OTC) として展開することができる。OTC は、2 つ以上のフォントを含む単一のフォントリソース (フォントファイル) であり、その主な利点は、フォント間でテーブルを共有した結果

として生じるフットプリント(容量)の削減である。源ノ角ゴシックは、OTCを使用した最初の主要なフォントの実装である。ウエイトごとに7つのOTCフォントがあるが、各OTCフォントは(ウエイトにより変化するが)5~10のフォントを含み、同じ「CFF」(コンパクトフォントフォーマット)テーブルを共有する。レギュラーとボールドウエイトには追加で、ASCII範囲が半角グリフにマップされている半角フォントが含まれる。さらに「スーパーOTC」と呼ばれるファミリーOTCは、45のフォントすべてとウエイトごとに7つの「CFF」テーブルを含んでいる。45の各フォントのサイズは16~18.5メガバイトの範囲であり、合計すると約750メガバイトである。ウエイトごとの7つの各OTCフォントのサイズは18~20メガバイトで、合計すると約136メガバイトである。45のフォントをすべて含むスーパーOTCフォントのサイズは123メガバイトである。このように、OTCフォントフォーマットでの実装は「CFF」や「sfnt」テーブルを共有することにより多大な容量の節約を可能にしている。ただし残る課題は、OTCをサポートする環境が十分でないという点であった。

最初に、Mac OS X 10.8 (2012年リリースのMountain Lion)とAdobe CS6(とその後のバージョン)のアプリケーションでOTCのサポートが開始され、その後マイクロソフトがWindows 10 Anniversary Update(バージョン1607、2016年8月2日にリリース)で初めてサポートを開始した。その後、さらに多くのフォントを含む「Mega」OTC(別々のSource Han OTCおよびNoto CJK OTC)と「Ultra」OTC(Source HanとNoto CJKとを合体した1つのOTC)と呼ばれるものを作成しリリースした。

Pan-CJK書体開発の最も興味深くかつ時間のかか

る側面は、各地域をサポートするためのグリフがどの文字に対してどの程度(いくつ)必要であるかという点である。大多数の漢字は1つのグリフのみ必要とするが、次いで多くの漢字が2つのグリフを必要とし、3つ以上のグリフを必要とする漢字の数は順次少なくなっていく。任意の文字に必要なグリフの数を特定するために、各地域の文字セット標準が参照された。このプロセスの過程で分かったことだが、文字セット標準というものが書体開発には必須である一方、さまざまな(文書化されていたりされていなかったりする)誤りが多くあり、必ずしも完全には信頼できないという事実がある。図-2では、源ノ角ゴシックバージョン2.001のフォント中の44,808コードポイントの文字について、Unicodeのコード範囲ごとに必要となるグリフ数(1~5)の文字がいくつあるかを示している。

縦書き用グリフのためにemボックス内での位置を変更する必要がある小仮名を除いて、従来の日本語フォントでは仮名のデフォルトグリフは縦書きと横書きで共有されている。一部の従来の日本語フォントには、横書きと縦書きのレイアウトに合わせてカスタマイズされた完全に別々な仮名のグリフセットが含まれることもあるが、デフォルトのグリフではないので、GSUB機能である「hkna」(Horizontal Kana Alternates)または「vkna」(Vertical Kana Alternates)で有効にする必要がある。Source Han書体は、横書きと縦書きの仮名グリフがデフォルトで別々に含まれている。縦書きという観点で述べると、Source HanフォントはUAX #50(Unicode Vertical Text Layout)に準拠した最初の実装の1つである。図-3は、源ノ角ゴシックの3つのひらがな文字の横書き用(青)と縦書き用(赤)の字形を示している。

		CJK Unified Ideograph Extensions							
		URO	A	B	C	D	E	F	Other
Glyphs	1	8,783	6,123	2,080	47	34	112	5	14,704
	2	7,642	447	28					191
	3	3,737	12						48
	4	746							1
	5	68							

■ 図-2
Unicodeの各ブロックにおける、必要なグリフの数

従来の日中韓フォントには通常、ギリシャ文字とキリル文字のグリフが含まれているが、ほとんどの場合全角グリフとして実装されている。Source Han 書体のギリシャ文字とキリル文字のグリフはプロポーショナルなため、従来のCJK フォントのグリフよりも実用性が高い。

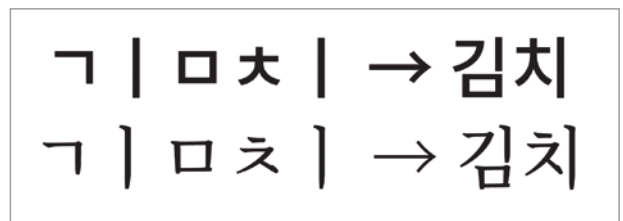
ハンゲルのサポートでは、組み合わせた字母 (jamo) のサポートを追加するという興味深いチャレンジをすることとなった。ハンゲル音節文字は、2つの字母からなる音節 399 文字、3つの字母からなる音節 10,773 文字の合計 11,172 文字で構成されており、そのうち日常で頻繁に使われるのは 3,000 文字以下とされている。現代における Jeju (제주말 / 濟州語) などの方言や、非現代 (古語) のハンゲルを表現するためには、現代では使われていない字母から構成される音節も必要とされる。2つまたは3つの字母からなるごく少数の現代では使われない音節文字は、「ccmp」(Glyph Composition/Decomposition) GSUB 機能を通してアクセス可能な事前合成グリフによりサポートされているが、すべての可能なシーケンス (合成音節文字) の数は 1,638,750 となる (現代音節に対応する 11,172 のシーケンスが除外された場合は 1,627,578)。Source Han 書体では字母を事前合成したグリフを実装しており、6セットの「初声子音」jamo (L)、2セットの「母音」jamo (V)、および4セットの「終声子音」jamo (T) が含まれている。「ljmo」(Leading Jamo Forms)、「vjmo」(Vowel Jamo Forms)、および「tjmo」(Trailing Jamo Forms) などの GSUB 機能は、適切な字母合成による音節文字の作成に有効である。図-4 は、例として Source Han フォントを使用して、韓国語の「キムチ」(김치) の3つおよび2つの字母の並びとそれに対応する音節を比較したものである。



■ 図-3 文字デザインにおける縦書き (赤) と横書き (青) の差異

Unicode の新しいバージョンには Source Han 書体の対象となる新しい文字が含まれることが多いため、Source Han の2つのファミリーは、一方のファミリーが他方に先駆けて特定文字のサポートやデザインの改良を実装することで他方のファミリーを飛び越してしまう傾向がある。たとえば源ノ角ゴシックバージョン 2.000 は、bopomofo (ㄅㄆㄇ / 注音) のグリフとレイアウトサポートを改善し、また繁体字中国語 (台湾と香港特別行政区) 用に、讠 (部首番号 #162) コンポーネントのデザインも改善した。源ノ明朝バージョン 2.000 でもこの点を同様に扱う予定だが、さらに新しい機能拡張が導入される可能性もある。今後ファミリーが成熟していくにつれ、この傾向は収束していくと期待する。

普段あまりフォントに興味を持たない人にとっても Source Han 書体をより興味深いものとする話題として、Source Han Serif バージョン 1.000 に含まれる複雑な漢字、その画数が 60 近くある「ピアン」(biáng) と呼ばれる、中国で麺類を表す文字の話がある。その文字は、簡体字中国語の形状であっても 40 画以上である。源ノ角ゴシックバージョン 2.000 は、日本のレストランチェーンの名称にも使われている、「たいと」または「おとど」と読む画数 84 画のさらに複雑な漢字をサポートすることで、源ノ明朝書体をリープフログ (馬跳び) した。これらはCJK 統合漢字拡張 G に含まれる本物の漢字ではあるが、第3面のコードポイントはまだ安定していないため、IDS (Ideographic Description Sequences) と「ccmp」GSUB 機能を使用してグリフにアクセスする擬似的な符号化形式と見なされる。図-5 は、源ノ角ゴシックバージョン 2.000 でサポートされている拡張 G の漢字と、それらの IDS および地域固有の文字の形を示している。



■ 図-4 ハンゲル字母の並びとそれに対応する音節

現在の我々の課題は、バリエブルフォント版の Source Han 書体、つまり5つのバリエブルフォントを内蔵するフォントコレクションを開発することである。各種ウエイトなどバリエーションごとに別のフォントリソースを必要とする従来のフォントとは異なり、バリエブルフォントはマスタデザイン間の補間技術を使用してウエイトや幅など任意のバリエーションを動的に生成することができる。ある書体が多くウエイトをそれぞれ別個のフォントとして持つ場合に比べ、最も細いウエイトと最も太いウエイトのみをマスタデザインとして保持するバリエブルフォント形式のフォント容量は、かなり小さくすることができる。バリエブルフォント版の Source Han 書体を作成しても、それを5つの言語固有のフォントとして利用することは、容量の縮小というバリエブルフォントの1つの利点を打ち消してしまうことになる。この点におけるチャレンジは、基盤（フォントの使用環境）の拡充にある。バリエブルフォント自身は一部の主要 OS やアプリケーションでサポートが開始されているが、バリエブルフォントコレクションのサポートは現時点でまだ開始されていない。

Source Han 書体についてのさらに詳細な、下記の（作成には大変骨が折れた）詳細な ReadMe ファイルを読むことを強くお勧めする。

源ノ角ゴシックの ReadMe : <https://github.com/adobe-fonts/source-han-sans/raw/release/Source-HanSansReadMe.pdf>

IRG	CN	TW	HK	JP	KR	IDS
0363 (UTC-0120)	KR	KR	KR	KR	恣	☞ 思
0418 (UTC-0112)		CN	CN	CN	CN	 ☞ 六 月 心 长 言 马 么 长 心
0419 (UTC-0079)		HK				 ☞ 六 月 心 长 言 马 么 长 心
0475 (UK-0260)					JP	 ☞ 雲 龍 雲 龍 雲 龍 龍

■ 図-5 拡張 G の漢字

源ノ明朝の ReadMe : <https://github.com/adobe-fonts/source-han-serif/raw/release/Source-HanSerifReadMe.pdf>

情報処理基盤へのメリット

Source Han 書体の目的の1つは、サポートする各言語に対して従来の地域固有のフォントと同様に動作することであるが、加えて数字の表現を欧文と日本語で切り替えたり、欧文のスマートクオートを日本語の引用符（開き括弧と閉じ括弧）に切り替えることができる「loc」GSUB 機能サポートなどの追加機能も内蔵している。つまり、ユーザが設定言語を日本語に指定した場合には従来の日本語フォントと同様に振る舞えるということであり、この場合フォントが持つ日本語以外の情報は無視される。これまで、特定の地域（たとえば「日本」など）に必要なグリフのみを含んだ地域固有のサブセットフォントを開発・配布してきたが、可能であれば複数言語をサポートする Pan-CJK バージョンの使用を強く推奨する。この場合、それぞれ任意のデフォルト言語を設定したり、その機能をサポートするアプリケーションで言語タグを指定したりすることで、Pan-CJK バージョンのフォントが有効に活用可能となる。

Source Han 書体の日本語に関するもう1つの目標は、すべての Adobe-Japan1-7 漢字、および JIS X 0208, JIS X 0212, および JIS X 0213 規格のすべての漢字を含めることであった。つまり、言語を日本語設定した場合には、すべての Adobe-Japan1 IVS に加え、フォーマット 14 (Unicode Variation Sequence) の「cmap」サブテーブルにおける必要最小限の数の SVS にアクセス可能となるため、既存の日本語フォントとの互換性が向上した。

前述したように、Source Han グリフセットはその独特な性質のため特殊用途の Adobe-Identity-0 ROS を使用する必要があった。これは、書体のグリフセットがアドビの公開 ROS（たとえば、日本語フォントならば Adobe-Japan1-7 など）に適合しない場合に指定

する ROS である。2つの Source Han 書体どちらも Adobe-Identity-0 ROS を使用しているが、Pan-CJK 書体の重要な特徴のため両者のグリフセットは必然的に異なっている。グリフの共有範囲をどう指定するかは、セリフ対サンセリフの違いのみならず、実際の書体デザインによっても変化するものである。つまり、両方の書体ともサポート対象の言語間でグリフを共有しているのだが、その分布はそれぞれ異なるのである。我々が開発した Adobe-Identity-0 フォントが、ほかのフォント開発者の参考となり、有意義な知見を与えるものとなることを願っている。最初の Adobe-Identity-0 フォントは、Adobe が 2009 年にリリースした「かづらぎ」である。また、Source Han 書体に続いて、Adobe が 2017 年末にリリースした「韶明朝」も、Adobe-Identity-0 ROS を利用している。

必ずしも意図したことではなかったのだが、現在 Source Han 書体は複数のフォーマットで配布されている。これには特定の環境における制限を回避するという理由もあったが、結果的にほかの書体にも役立つ新しい技術の可能性を探る機会を得ることとなった。フォーマットの 1つは OTC (OpenType Collection) と呼ばれ、1つのフォントリソース中に複数のフォントを含み、それらがキーとなる「sfnt」テーブルを共有するものである。アドビの韶明朝書体は、単一の「CFF」テーブルを参照する 4つのフォントを内蔵した OTC フォントとして、OTC フォーマットの利点を踏襲している。現在進行中の挑戦として、バリエブルフォント版の Source Han 書体ファミリーに対する新たな展開フォーマットとして、バリエブルフォントコレクションの実装を実験している。この目的のために、我々は 2019 年 1 月末に Source Han 書体をシミュレートすることを目的とした一連のテストフォントを、6 および 12 のフォントを内蔵するバリエブルフォントコレクションとしてリリースした。これは、将来バリエブルフォントコレクションがリリースされる際、少なくとも主要な環境下で動作がサポートされることを望んでのことである。

見落としがちな点だが、Source Han 書体の利点と

して、フォント開発ツールはもとよりフォントを使用する環境の側の誤った憶測に基づく問題を明らかにする上で、非常に役立つことが挙げられる。これにより、今後同様のフォントがリリースされ使用された場合に、意図した通りに機能する方向へと導くことが可能となる。その種の誤った憶測があり得る例としては、フォント内のグリフの数が 65,535 のアーキテクチャ上の上限値に達している点についてのもの、OTC に複数フォントを内蔵している点、またこれまでにない行の高さや行間の決定方法などが含まれる。

本稿の前半で述べたように、本来の Pan-CJK 書体ファミリーは東アジアの言語間で一貫した書体デザインを提供することができるが、同時に漢字や句読点に関して地域固有の慣例を尊重している。このようなフォントを用いることで、企業は一貫性のある体裁とイメージを持った多言語のドキュメント、販促資料、およびその他の付随資料を作成することができるようになる。このことはブランディング上の種々の目的において重要な意味を持つ。たとえば、日本の公共交通システムにおける多言語の標識などが代表的な例として挙げられる。

最後に、Source Han フォントはオープンソースとして無料で使用できるため、その使用に対する障壁を事実上排除する。もちろん、高品質の Pan-CJK フォントをオープンソースとして利用可能にするということは、一見既存の商用フォントライセンスビジネスを混乱させるものと映るかもしれない。しかし、現実には、そのような高品質のオープンソース Pan-CJK フォントの数は限定されたものにならざるを得ないのである。なぜなら、その開発には相当な時間と技術と努力が必要とされるからである。

(2019 年 7 月 13 日受付)

■ Ken Lunde lunde@adobe.com

Adobe にて日中韓越の漢字圏フォント開発を担当。CJKV 日中韓越情報処理 (オライリー出版) の著者。2018 年に Unicode ブルドッグ賞を受賞。同年に Unicode テクニカルディレクターに就任。

用語集

- **Adobe-Japan1** - 現在市場に出回っている数百または数千の日本語フォントの基盤となる Adobe の公開日本語グリフセット。現在の Supplement は 7 (Adobe-Japan1-7) で、23,060 のグリフ (CID 0 から 23059) を指定している。
- **BMP** - 基本多言語面。文字が割り当てられた最初の Unicode 面で、第 0 面とも呼ばれる。
- **CFF** - コンパクトフォントフォーマット。PostScript ベースの、名前キー方式および CID キー方式フォントのコンパクトな表現。
- **CID** - 文字 ID。CID キー方式フォントのグリフに割り当てられる一意の整数値。
- **CID キー方式フォント** - グリフと CID を関連付け、CMap リソースを使用して文字コードを CID に対応づける PostScript ベースのフォントフォーマット。
- **CMap/cmap** - Unicode などの文字コードをグリフにマップするリソース (CMap) または [sfnt] テーブル (cmap)。
- **コードポイント** - 文字に対応する一意の数値。
- **フォントフォールバック** - 選択したフォントにグリフが含まれていない場合に、使用可能なフォントリソースを使用して文字をレンダリングする手法。
- **コードポイントとシーケンス** - 「字」=U+5B57 など、ある文字は単一のコードポイントとして表すことができる。一方、「あ」= U+3042「あ」と U+3099 のように、1 文字が一連の文字の列 (シーケンス) で構成されることもある。
- **Em-box** - 高さと同幅が文字 [M] の幅にほぼ対応する長方形のスペース [日本語では全角に相当する]。
- **フォントコレクション** - 1 つ以上の [sfnt] テーブルを共有する 2 つ以上のフォントを含めるためのコンテナフォーマット。
- **フォントファミリー** - 一様なデザインの 1 つ以上のフォントのセット。
- **GSUB 機能** - Glyph SUBstitution, グリフ置換。代替グリフをグリフに置き換える OpenType 機能。
- **IDC** - 漢字構成記述文字。より大きなコンポーネントを形成する 2 つまたは 3 つのコンポーネントの位置を説明する文字。
- **IDS** - 漢字構成記述文字列。漢字を視覚的に記述するために 1 つ以上の IDC とそのコンポーネントを使用する文字列。
- **Jamo** - 子音または母音に対応し、ハングル音節を形成するために 2 つまたは 3 つのジャモシーケンスでクラスタ化されている文字。字母。
- **言語タグ** - 言語を文字、段落、または文書レベルで文書のテキストに関連付けることを可能にし、選択されたフォントの機能に応じてテキストの表示方法が異なる場合がある、アプリケーションレベルの機能。
- **Multiple-em character** - グリフが一般的な文字のグリフよりも 2 全角以上広い、または高い文字。例としては、2- および 3-em ダッシュがある。
- **OpenType/CFF フォント** - グリフが OpenType の [CFF] テーブルに含まれている、最も広範に使用されている PostScript ベースのフォントフォーマット。
- **第 3 面** - 文字が割り当てられる次の Unicode 面。
- **sfnt テーブル** - スケーラブルフォント。オープンタイプのような最新技術に基づいたフォントのリソースに含まれるテーブルのタイプ。
- **疑似エンコーディング** - エンコードされていない漢字を表すために IDS を使用するなど、文字を表現するための代替方法。
- **サンセリフ** - セリフがなく、水平方向と垂直方向の画線が均一の太さである書体スタイル。日本語では「ゴシック」と呼ばれる。
- **セリフ** - 大きい画線の最後に小さい画線を含み、水平方向と垂直方向の画線の太さの違いの大きな書体様式。日本語では「明朝」と呼ばれる。
- **SVS** - Standardized Variational Sequence, 基底文字とそれに続く固有の視覚的形態に割り当てられる字形選択子からなる Unicode シーケンス。
- **TrueType フォント** - グリフが OpenType の [glyf] テーブルに含まれている一般的なフォントフォーマット。
- **書体** - 共通の様式またはデザイン上の特性を示すグリフを含むフォントを包含する集合。
- **書体デザイン** - 書体様式から独立した、書体の個別具体的なデザイン上の特性。
- **書体様式** - 複数の書体に共通して顕著に現れる共通のデザイン上の特性。一般的な書体様式として、セリフ (明朝体) とサンセリフ (ゴシック体) がある。
- **バリエアブルフォント** - マスタデザイン間の補間を使用して、太さや幅などの特定のバリエーションを動的に生成できるフォント。
- **縦書きグリフ** - 小仮名、大括弧、句読点など、縦書きでのみ使用される文字の形状。
- **垂直メトリック** - 行の高さと行間に関連付けられているフォントのメトリック設定。