

Repertuar znaków fontów komputerowych

Andrzej Tomaszewski

Fonty komputerowe są programami zawierającymi informacje o graficznych kształtach liter, dane typometryczne; mają też własny repertuar znaków danego kroju pisma. Są zatem **cyfrowymi nośnikami pism**, które całkowicie wyparły z przemysłu poligraficznego klasyczne nośniki pism drukarskich, takie jak czcionki, matryce lino- i monotypowe albo fotomatryce. Klasycznych nośników używa się dzisiaj już tylko w skansenach drukarstwa – placówkach typu muzealnego albo amatorskich oficynach bibliofilskich.

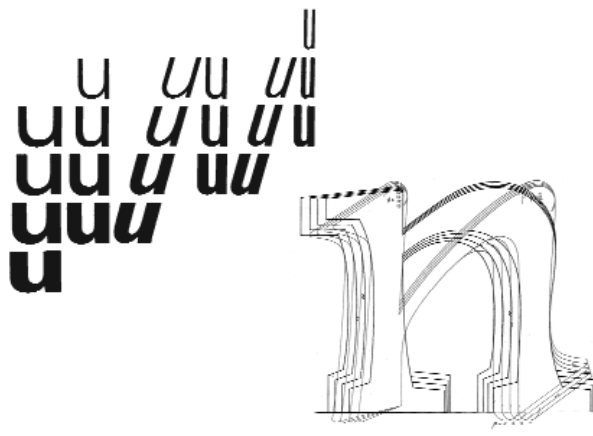
Waga i znaczenie problematyki związanej z produkcją fontów oraz ich walorami użytkowymi często są lekceważone w naszym środowisku poligraficznym, jakkolwiek większość z nas zdaje sobie sprawę, że pismo jest fundamentem wszelkiej drukowanej publikacji, jej zawartości treściowej, a także jakości informacyjnej i estetycznej.

Właściwości graficzne fontu konkretnego kroju pisma możemy nazwać **atrybutami graficznymi fontu**. Składają się one na charakterystyczny obraz wizualny pisma drukarskiego

ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890
ABCDEF abcdefghijklmnopqrst 1234567890

Przykład rodziny kroju pisma.

W przypadku kroju Minion Pro przynajmniej 8 fontów dla podstawowych odmian



Principia projektowania rodziny kroju pisma

o jednolitych cechach graficznych: stylu, rytmie, proporcjach, dukcie, właściwościach optycznych itp. Krój pisma może mieć wiele odmian (rodzina kroju pisma), często dość znacznie różniących się między sobą, ale zawsze związanych wspólnymi elementami graficzno-morfologicznymi. Krój pisma jest dziełem autorskim podlegającym ochronie prawnej.

Właściwości techniczne, takie jak: określenie pola znaków i bazowej linii pisma, hinty, wartości kerningu albo tak zwane funkcje zecerskie nazwijmy **atrybutami technicznymi fontu**. Służą one prawidłowej pracy zecerskiej, a więc technice składania i obliczania typograficznego układu znaków, wierszy tekstu, łamów, kolumn i wreszcie całej publikacji. W tej grupie atrybutów są również informacje dla konstrukcji pisma mniej istotne, na przykład o nazwie rodziny kroju pisma i jego odmianie, nazwie producenta albo też posiadaczu praw autorskich.

Trzeci rodzaj właściwości fontów to ich format, rozumiany jako programistyczna struktura danych. Te właściwości można nazwać **atrybutami programistycznymi fontu**. Należy do nich na przykład sposób zapisu kształtów liter za pomocą krzywych matematycznych, najczęściej krzywych Béziera trzeciego stopnia lub b-skle-

Do składania tekstów [Σαλαμινος] wielojęzycznych potrzebne są fonty, których forma wizualna [Санкт-Петербург] jest jak najbardziej do siebie zbliżona [וּפִי-בִּיבְא־לֵת].

Do składania tekstów [Σαλαμινος] wielojęzycznych potrzebne są fonty, których forma wizualna [Санкт-Петербург] jest jak najbardziej do siebie zbliżone [וּפִי-בִּיבְא־לֵת] .

Σαλαμινος Σαλαμινος
Санкт-Петербург Санкт-Петербург
וּפִי-בִּיבְא־לֵת וּפִי-בִּיבְא־לֵת

Potrzeba projektowania spójnych graficznie alfabetów dla pism dzielowych

nych do składania długich tekstów. Projektowaniem pism tytułowych oraz akcydensowych rządzą inne zasady związane z ich funkcjami użytkowymi.

Od kilku lat coraz szerzej wchodzi do praktyki DTP i poligraficznej, a w zasadzie szerzej – informatycznej, fonty oparte na tak zwanym unikodzie. Unicode jest to sposób kodowania znaków, który teoretycznie obejmuje wszystkie pisma używane na naszym globie. Unikod normalizują dwa standardy: Unicode oraz norma ISO 10646. Zestawy glifów (znaków) w obu systemach są identyczne, istnieją jednak między nimi drobne różnice wynikające niuansów zecerskich. Unicode jest rozwijany przez międzynarodowe konsorcjum największych firm komputerowych i producentów oprogramowania. Konsorcjum współpracuje z organizacją ISO oraz instytucjami naukowymi i członkami tak zwanych grup użytkowników z wielu krajów, w tym również użytkowników polskich. Do niedawna mieliśmy do czynienia z kilkuset standardami kodowania używanymi na świecie do różnych celów: lingwistycznych i semiograficznych. Przykładowo, dla języków posługujących



Umieszczanie grup glifów w fontach unikodowych

Latin Extended Additional
Range: 1E60–1EFF

This file contains an excerpt from the character code tables and list of character names from The Unicode Standard, Version 5.2.

This file may be changed at any time without notice to reflect errata or other updates to the Unicode Standard. See <http://www.unicode.org/errata> for an up-to-date list of errata.

See <http://www.unicode.org/char/1E60-1EFF> for access to a complete list of the latest character code charts. See <http://www.unicode.org/Public/5.2.0/UnicodeList.html> for a complete list of the latest character code charts for Unicode 5.2.

Disclaimer

These charts are provided as the online reference to the character contents of the Unicode Standard, Version 5.2, but do not provide all the information needed to fully represent individual codes using the Unicode Standard. For a complete understanding of the use of the characters contained in this file, please consult the appropriate section of the Unicode Standard, Version 5.2, online at <http://www.unicode.org/unicode/5.2/>, as well as Unicode Standard Annexes #9, #11, #14, #15, #24, #25, #26, #27, #28, #29, #30, #31, #32, and #33, and the other Unicode Technical Reports and Standards, and the Unicode Character Database, which are available online.

See <http://www.unicode.org/char/1E60-1EFF> for a list.

A thorough understanding of the information contained in these additional sources is required for a successful implementation.

Facts

The shapes of the reference glyphs used in these code charts are not prescriptive. Considerable variation is to be expected in actual form. The particular forms used in these charts were provided to the Unicode Consortium by a number of different font designers, who own the rights to the fonts.

See <http://www.unicode.org/char/1E60-1EFF> for a list.

Terms of Use

You may freely use these code charts for personal or internal business use only. You may not incorporate them either whole or in part into any product or publication, or otherwise distribute them without express written permission from the Unicode Consortium. You may provide links to these charts.

The fonts and font data used in production of these code charts may NOT be extracted, or used in any other way in any product or publication, without permission or license granted by the typeface owner(s).

The Unicode Consortium is not liable for errors or omissions in this file or the standard itself. Information on characters added to the Unicode Standard since the publication of the most recent version of the Unicode Standard, as well as on characters currently being considered for addition to the Unicode Standard can be found on the Unicode web site.

See <http://www.unicode.org/unicode/5.2/> for a list.

Copyright © 1991–2009 Unicode, Inc. All rights reserved.

3040

Hiragana

	304	305	306	307	308	309
0	く	だ	ば	む	み	
1	あ	け	ち	ば	め	ゑ
2	あ	げ	ち	ひ	も	を
3	い	こ	つ	び	や	ん
4	い	こ	つ	び	や	ず
5	う	さ	つ	ふ	ゆ	か
6	う	さ	て	ぶ	ゆ	け
7	え	し	で	ぶ	よ	
8	え	じ	と	へ	よ	
9	お	す	ど	べ	ら	ゝ
A	お	ず	な	べ	り	ゝ
B	か	せ	に	ほ	る	
C	が	ぜ	ぬ	ぼ	れ	ゝ
D	き	ぞ	ね	ぼ	ろ	ゝ
E	ぎ	ぞ	の	ま	わ	ゝ
F	く	た	は	み	わ	ゝ

300

The Unicode Standard 5.2. Copyright © 1991–2009 Unicode, Inc. All rights reserved.

13A0

Cherokee

	13A	13B	13C	13D	13E	13F
0	D	I	G	J	V	B
1	R	Γ	Λ	Ω	ϥ	Ꭰ
2	T	Ꭱ	h	R	P	Ꭲ
3	Ꭳ	W	Z	L	G	Ꭴ
4	Ꭵ	Ꭶ	q	w	v	Ꭷ
5	I	Ꭸ	Ꭹ	S	Ꭺ	Ꭻ
6	S	G	T	Ꭼ	K	Ꭽ
7	Ꭾ	M	Ꭿ	Ꮀ	Ꮁ	Ꮂ
8	Ꮃ	q	Ꮄ	Ꮅ	Ꮆ	Ꮇ
9	Y	Ꮉ	v	V	G	Ꮊ
A	Ꮋ	Ꮌ	Ꮍ	Ꮎ	Ꮏ	Ꮐ
B	J	H	Ꮏ	Ꮑ	Ꮒ	Ꮓ
C	E	Ꮎ	Ꮏ	Ꮔ	Ꮕ	Ꮖ
D	Ꮏ	Ꮏ	Ꮏ	Ꮗ	Ꮘ	Ꮙ
E	P	Ꮏ	Ꮏ	Ꮚ	L	G
F	Ꮏ	Ꮏ	Ꮏ	Ꮜ	C	Ꮝ

300

The Unicode Standard 5.2. Copyright © 1991–2009 Unicode, Inc. All rights reserved.

2D80

Ethiopic Extended

	2D8	2D9	2DA	2DB	2DC	2DD
0	ᐃ	ᐄ	ᐅ	ᐆ	ᐇ	ᐈ
1	ᐉ	ᐊ	ᐋ	ᐌ	ᐍ	ᐎ
2	ᐏ	ᐐ	ᐑ	ᐒ	ᐓ	ᐔ
3	ᐕ	ᐖ	ᐗ	ᐘ	ᐙ	ᐚ
4	ᐛ	ᐜ	ᐝ	ᐞ	ᐟ	ᐠ
5	ᐡ	ᐢ	ᐣ	ᐤ	ᐥ	ᐦ
6	ᐧ	ᐨ	ᐩ	ᐪ	ᐫ	ᐬ
7	ᐭ	ᐮ	ᐯ	ᐰ	ᐱ	ᐲ
8	ᐳ	ᐴ	ᐵ	ᐶ	ᐷ	ᐸ
9	ᐹ	ᐺ	ᐻ	ᐼ	ᐽ	ᐾ
A	ᐿ	ᑀ	ᑁ	ᑂ	ᑃ	ᑄ
B	ᑅ	ᑆ	ᑇ	ᑈ	ᑉ	ᑊ
C	ᑋ	ᑌ	ᑍ	ᑎ	ᑏ	ᑐ
D	ᑑ	ᑒ	ᑓ	ᑔ	ᑕ	ᑖ
E	ᑗ	ᑘ	ᑙ	ᑚ	ᑛ	ᑜ
F	ᑝ	ᑞ	ᑟ	ᑐ	ᑠ	ᑡ

300

The Unicode Standard 5.2. Copyright © 1991–2009 Unicode, Inc. All rights reserved.

Przykłady tablic Unicode

się alfabetem łacińskim mieliśmy dziesiątki kodowań opartych na amerykańskim standardzie ASCII, na przykład w naszym kraju Latin 2, Mazovia, Apple CE, PL-Euro itp. Wiele fontów z tymi historycznymi niejako systemami kodowania jest jeszcze w użyciu w pracowniach DTP (fonty w rodzaju PostScript-Type1 oraz TrueType), bowiem z racji matematycznie zdefiniowanych kształtów znaków nie sprawiają one kłopotu w nowoczesnych naświetlarkach filmów i form drukowych. Są jednak oczywiście kłopotliwe w procesach przetwarzania danych – trzeba wówczas stosować specjalnie napisane konwertery kodów. Pamiętajmy, że istnieją niezbadanej wielkości archiwa zapisane w rozmaitych

systemach kodowania. Na przykład wznowienie książki naukowej sprzed dziesięciu lat, zachowanej w postaci plików dowolnego programu edytorskiego nie jest dzisiaj przedsięwzięciem banalnym – jest pracochłonne i wymaga specyficznej wiedzy.

W dzisiejszych fontach o strukturze unikodowej mamy przewidziane miejsce na wszystkie języki świata. Pamiętajmy jednak, że fonty robi się dla konkretnego kroju pisma lub jego odmiany, na przykład Adobe Garamond Semibold, Helvetica Oblique, Antykwa Półtawskiego Italic itp. Z tego wynika, że żaden font nie ma kompletnego zestawu znaków, choć jego struktura programistyczna jest przygotowana na dodanie

ABCDEFGH 1234567890
 abcdefghijklm
 1234567890
 .,:«»!&fiflß
 aenrrttQ
 abem₂₃₄₅₂₃

ABCDEFGH 1234567890
 abcdefghijklm
 1234567890
 .,:«»!&fiflß
 ABCDEF
 GHJML
 1234567890
 &ffiffi fltst

ABCDEFGH 1234567890
 abcdefghijklm
 1234567890
 .,:«»!&fiflæœ
 fflffABCDEFG
 1234567890
 !& itr₂₃₄₅₆₇₈₉
 ~~~~~~  
 ~~~~~~  
 ~~~~~~

Różne litery alternatywne, ligatury, indeksy, kapitaliki i inne formy graficzne umieszczane w fontach jako elementy konkretnych krojów pisma

na przykład pisma ormiańskiego, tajskiego czy Indian Czirokezów.

Istnieje dzisiaj jeszcze sporo fontów, które mają znaki nie przewidziane przez standardy Unicode i takich, które wyprodukowano do bardzo specyficznych zastosowań. Są wśród nich fonty z semiogramami, czyli różnymi znakami specjalnymi (np. turystycznymi, drogowymi, kartograficznymi itp.), ze znakami różnych notacji, na przykład muzycznych, matematycznych, chemicznych i innych, z ornamentami oraz elementami teł i poddruków, albo też fonty przygotowane do jednego, specjalnego przedsięwzięcia wydawniczego. Wszystkie one mają jakieś własne kodowanie, które w przypadku późniejszego przetwarzania wykonanego nimi składu wymaga stosowania specjalnie napisanych konwerterów kodu.

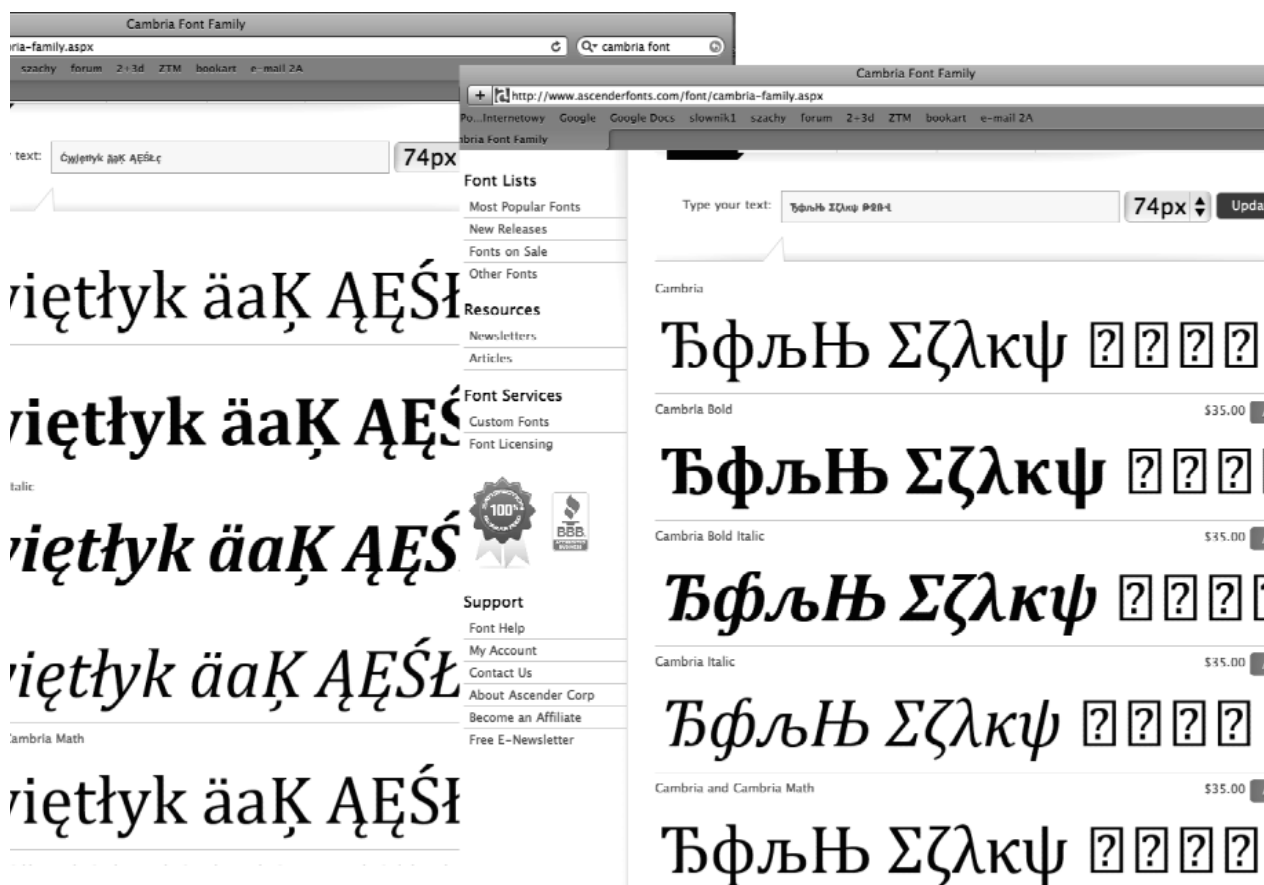
Współcześnie kroje pism projektowane do składania tekstów naukowych i szkolnych mają najczęściej kodowanie unikodowe, a ich projektanci starają się zapisać jak najwięcej miejsc zarezerwowanych na dane glify. Na przykład czeski

postscriptowy Skolar, zaprojektowany jako pismo podręczników, ma w odmianie podstawowej 1066 glifów, a truetypeowa Cambria firmy Microsoft ma 6988 glifów. Podobno najwięcej glifów ma popularny Arial firmy Monotype – ponad 50 tys. glifów w odmianie podstawowej. Dużą liczbę glifów mają fonty od lat używane do tekstów naukowych, na przykład Times New Roman oraz TeX-owy Computer Modern – nad rozbudowaniem repertuaru jego znaków od lat pracuje grono fontowych programistów z polskiej Grupy Użytkowników Systemu TeX.

Ściśle związaną z kodowaniem, a jednak osobną sprawą, są fonty OpenType, które wewnątrz własnego oprogramowania mają wmontowane tak zwane funkcje zecerskie. OpenType umożliwia swobodne zamiany glifów, na przykład liter tekstowych na kapitaliki, cyfr tabelarycznych na nautyczne, wymienianie liter alternatywnych, stosowanie ligatur ortograficznych i estetycznych oraz abrewiacji. Wszystkie te czynności wymagają, aby font miał w swoim repertuarze glify, których nie znajdziemy w uniko-dzie. Należą do nich na przykład kapitaliki czy indeksy górne i dolne, specjalnie zaprojektowane tak, aby nie wyglądały na zmniejszone majuskuły lub cyfry, ale by pasowały optycznie do podstawowego tekstu. Mają też wymienne warianty poszczególnych liter lub grup liter, albo też zestawy wariantowych cyfr.

W programach adobowskich, na przykład w InDesign i Illustrator CS4, owe funkcje zecerskie działają poprawnie. We wcześniejszych wersjach było kilka błędów, co powodowało, że pewne funkcje zecerskie przy stosowaniu niektórych fontów działały nieprawidłowo. Nie była to jednak kwestia fontów, tylko niekompletnej implementacji OpenType Layout w programach Adobe. Nawiasem mówiąc, jest ona nadal niekompletna, tylko mniej.

Problem ten nie dotyczy zresztą tylko Adobe. Programy praktycznie wszystkich dużych producentów oprogramowania używającego funkcji zecerskich OpenType mają mocno niekompletną implementację OpenType Layout. W programach Adobe nie działają jedne aspekty, zaś w programach Microsoft drugie. W programach Apple inne, a jeszcze inne w aplikacjach pod marką QuarkXPress. Niestety więc realia są



Przy wyborze fontów do zakupu w internecie istnieją możliwości sprawdzenia repertuaru znaków

takie, że fonty OpenType – szczególnie z bardzo rozbudowanym repertuarem znaków i skomplikowane – trzeba nabywać pod kątem konkretnego oprogramowania.

Sytuacja zmienia się powoli, ale wydaje się, że w dobrym kierunku. Na przykład w „zachodnich” wersjach InDesign CS4 po raz pierwszy można używać pism arabskich czy hebrajskich (nie jest to jeszcze proste, ma się poprawić w następnych wersjach), w Microsoft Office 2010 (wersje beta dostępne publicznie) działają funkcje zecerskie cyfr nautycznych czy też warianty kontekstowe w alfabetach europejskich. W programach Apple dla systemu Mac OSX 10.6 działają na przykład warianty kontekstowe, które nie działały w Mac OSX 10.5.\*

Jak więc widzimy, jeszcze nieco wody upłynie w Warcie, zanim wszystko zadziała bez zarzutu,

\* Wybrane aspekty działania funkcji zecerskich według analizy wykonanej przez Adama Twardocha z firmy FontLab.

a zwłaszcza zanim graficy dorysują brakujące glify, choćby tylko do najpopularniejszych krojów pisma.

Zestawy znaków poszczególnych alfabetów rozumianych jako systemy zapisu możemy roboczo nazwać zestawami prymarnymi. Mniej więcej odpowiadają one temu, co się wiąże z ortografią danego języka. W przypadku alfabetu łacińskiego mamy tu majuskuły, czyli wielkie litery, oraz minuskuły, czyli litery małe (tzw. tekstowe). Prymarny zestaw liter naszego alfabetu został jednak uzupełniony wieloma znakami o kształtach wynikających z potrzeb praktycznych albo estetycznych. Te formy literne powstały w ciągu całej historii piśmiennictwa i drukarstwa, a także powstają dzisiaj. Trzeba zatem znaleźć dla nich miejsce w repertuariach fontów komputerowych. Należą do nich na przykład uncjalne lub kapitalikowe formy liter, warianty poszczególnych liter składane na początku, w środku lub na końcu tekstu, następnie ligatury, czyli połączenia dwóch lub więcej liter



