

Standard Unicode w typografii

Wprowadzenie do standardu Unicode

Unicode – według formalnej definicji – jest standardem kodowania znaków zaprojektowanym do wymiany, przetwarzania i wyświetlania dokumentów pisanych w każdym języku i dotyczących każdej dziedziny współczesnego świata, w tym również tekstów w językach wymarłych i historycznych. Standard ten został stworzony przez naukowców, informatyków i lingwistów z wielu firm branży komputerowej i instytucji naukowych, zrzeszonych w „Konsorcjum Unicode” – niekomercyjnej organizacji, kierowanej przez przedstawicieli firm komputerowych [1]. Unicode jest standardem ogólnie dostępnym i w zasadzie bezpłatnym, choć niestety obwarowanym zastrzeżeniami praw autorskich i patentami.

Historia standardu rozpoczęła się w późnych latach osiemdziesiątych ubiegłego wieku od prac nad stworzeniem formatu fontu TrueType. Pierwsza wersja standardu Unicode ukazała się w 1991 r., dynamiczny rozwój i popularyzacja nastąpiły pod koniec lat dziewięćdziesiątych. Najnowsza, aktualna wersja standardu (6.2.0) została opublikowana we wrześniu 2012.

Jako ciekawostkę z prehistorii warto wspomnieć o standardzie Unicode opisanym już w 1889 r. w książce pod tytułem: „Unicode: The Universal Telegraphic Phrase-Book. a Code of Cypher Words for Commercial, Domestic, and Familiar Phrases in Ordinary Use in Inland and Foreign Telegrams” [2] – co można przetłumaczyć jako: „Unikod: Uniwersalna książka skrótów telegraficznych. Słowa kodujące zwroty handlowe, domowe i potoczne wykorzystywane powszechnie w telegramach krajowych i zagranicznych”.

Założenia i praktyka standardu Unicode

Założenia standardu Unicode obejmują: ogólnoświatowy (szeroki) zestaw znaków, ich logiczny porządek, efektywność przetwarzania dokumentów, ich jednolitość, stabilność standardu, zachowanie semantyki znaków i niezależność od ich wyglądu. Standard obejmuje także algorytmy łączenia znaków (np. litery akcentowane), ich dekompozycje, a także wersje znaków w zależności od kontekstu (co dotyczy głównie języków afroazjatyckich).

W specyfikacji standardu założenia te zostały opatrzone hasłami:

- uniwersalny repertuar;
- porządek logiczny;
- efektywność;
- ujednolicenie;
- znaki, a nie kształty;
- dynamiczne łączenie;

- znaczenie (semantyka);
- stabilność;
- zwykły tekst (bez instrukcji);
- przenośność.

Podsumowaniem i uzupełnieniem powyższych określeń jest zdanie: „Unicode nadaje każdemu znakowi niepowtarzalny numer, niezależny od komputera, niezależny od programu, niezależny od języka”.

Warto zauważyć, że już w samych założeniach kryją się pewne wewnętrzne sprzeczności. Założenie aktualności standardu i uwzględnienia w standardzie każdego znaku stoi w opozycji do postulatu stabilności. W praktyce obecnie zdecydowanie zwycięża stabilność, wspierana dodatkowo przez biurokratyczną bezwładność konsorcjum Unicode. Zmiany standardu odbywają się powoli i poprzedzone są kolejnymi wersjami projektów i raportów. Wersje standardu opublikowane w ciągu ostatnich 6 lat różnią się głównie w bardzo egzotycznych dla nas, Europejczyków, zakresach. Należy jednak pamiętać, że fundamentem pod obecną stabilność standardu są jego poprzednie wersje, obciążone wieloma błędami i zaszłościami z pierwszego okresu rozwoju standardu, a także niestety odziedziczonymi po wcześniejszych rozwiązaniach.

Drugim dylematem wynikającym z założeń jest konflikt pomiędzy efektywnością i logicznym porządkiem a kompletnością i rozległością zestawu znaków. Pierwsze wersje standardu zawierały ok. 30 tys. znaków – 5 tys. w obszarze podstawowym (BMP) i 25 tys. znaków wschodnioazjatyckich (CJK). Obecnie standard zawiera ponad 110 tys. znaków. Wprawdzie aktualnie standard przewiduje możliwość ponumerowania ponad miliona znaków, ale jest to wynikiem zweryfikowania pierwotnych założeń, które zostały zdecydowanie przekroczone. (W 1988 roku Joe Becker, jeden z twórców standardu, pisał: „[potrzebujemy] niewątpliwie znacznie mniej niż 16 tysięcy znaków” [3].)

Unicode dla czytelnika i pisarza

Przeciętnemu użytkownikowi komputera standard Unicode zapewnia wygodną obsługę tekstów w dowolnym języku, ich przenośność pomiędzy programami i systemami operacyjnymi, a także wymianę dokumentów z osobami z całego świata. Upowszechnienie tego standardu pozwala zapomnieć o koszmarze „standardów kodowań” (dla samego języka polskiego używaliśmy ponad dziesięć stron kodowych, w tym: ISO-8859-2, Windows 1250, Latin 852 i Mazovia).

Dzięki standardowi Unicode przenoszenie tekstów pomiędzy programami czy systemami operacyjnymi stało się łatwe i naturalne. Można skopiować fragment dokumentu w formacie PDF, wstawić go do automatycznego tłumacza w przeglądarce internetowej, a wynik tłumaczenia wysłać e-mailem osobie, która używa komputera z innym systemem operacyjnym – i nie zastanawiać się, jakiego kodowania używa który program. W tym zakresie standard Unicode sprawdza się dobrze. Choć oczywiście nadal zdarzają się „nieporozumienia” i czasami pojawiają się niepożądane symbole, ale wynika to najczęściej z dziedzictwa starych kodowań oraz programów nieobsługujących Unicode.

Także autorom książek i naukowcom (zwłaszcza lingwistom) Unicode oddaje nieocenione usługi. Cytaty w obcych językach (nawet archaicznych i wymarłych), zapisy fonetyczne, dodatkowe symbole akcentowe – to wszystko jest łatwo dostępne przy założeniu, że używany font zawiera potrzebne znaki – ale ta kwestia zostanie omówiona dalej.

Ostatnie wersje standardu zawierają także poszerzone wsparcie dla opisywania wzorów matematycznych. W połączeniu z nową technologią fontów o nazwie OpenType czyni to Unicode użytecznym również dla matematyków, fizyków i adeptów innych nauk ścisłych, zwłaszcza gdy chcą przygotować tekst do publikacji.

Unicode dla zecera

Osobom zawodowo związanym z książką, poczynając od redaktorów i korektorów po osoby zajmujące się składem, standard Unicode, choć nadal zdecydowanie przydatny, pokazuje też swoje ciemniejsze strony.

Zalety standardu omówione powyżej także w tej branży są bardzo istotne. Zwłaszcza przy wydawaniu publikacji wielojęzycznych, w egzotycznych z europejskiego punktu widzenia językach, a także czasopism z międzynarodowym gronem autorów. Wtedy uniwersalność i jednolitość kodowania ma decydujące znaczenia dla usprawnienia pracy i uniknięcia błędów wynikających np. z konwersji na inną stronę kodową.

Jednak już przy pracy redakcyjnej niektóre cechy standardu Unicode stają się jego wadami. Na przykład niezależność od języka powoduje, że litery ABC (czyli: U+0410 U+0412 U+0421) i inne są nieodróżnialne od liter ABC (czyli: U+0041 U+0042 U+0043 – por. rysunek 1), a ich podmiany zdarzają się całkiem często, np. w angielskojęzycznych tekstach autorów rosyjskich (dla jasności: pierwsza seria znaków to litery cyryliczne, a druga to litery łacińskie). Takie błędy, choć niewidoczne, wymagają korekty, inaczej ujawnią się na dalszym etapie prac, na przykład brakiem możliwości przeszukiwania tekstu (bo nie da się, szukając nazwiska zaczynającego się od C znaleźć tego, które zaczyna się od rosyjskiego S) czy błędnymi podcięciami (bo obecne fonty z zasady nie mają kerningu pomiędzy literami należącymi do różnych alfabetów).

BOA BOA BOA

B O A BOA BOA

ⓅⓄⓂ (b)(o)(a)

Rys. 1. Unicode'owe znaki z jednego fontu *DejaVu Serif*; pierwszy wiersz: łacińskie, cyrylica, greka; drugi: fullwidth form, matematyczne bezszeryfowe, indeksy górne i trzeci wiersz: kółeczkowane oraz nawiasowane

Na etapie prac redakcyjnych istotne staje się także to, że standard Unicode nie przenosi informacji o wyglądzie znaków. Dyrektywy dotyczące formatowania (np. wyróżnienia, cytaty, podtytuły) trzeba zapisywać w inny sposób: albo przechowując dokumenty w postaci powiązanej z jakimś programem (np. format .odt czy .doc), albo stosując jakiś język znaczników (np. XML czy \TeX). Oba te rozwiązania niestety niwelują część zalet standardu Unicode.

Także na dalszym etapie prac nad składem, a szczególnie w przypadku udostępniania dokumentu w postaci elektronicznej, występuje zamieszanie związane z niekonsekwentnym odcinaniem się standardu Unicode od formy znaku. Na przykład przy zastosowaniu w składzie dokumentu fontu zawierającego oprócz podstawowej także wersję kapitalikową (której Unicode nie uwzględnia) może się zdarzyć, że fragmentów tekstu złożonych tym wariantem fontu nie będzie można przeszuwać (choć odpowiednio „inteligentny” system DTP jest w stanie temu zapobiec). Kolejnym przykładem niekonsekwencji jest to, że standardowe ligatury drukarskie (typu: „fi”, czy „ff”) wchodzi w zakres standardu Unicode i w związku z tym wymagają specjalnego traktowania. Pojawiają się również niespodziewane formy znaków, na przykład znaki w nawiasach czy kółeczkach – por. rysunek 1.

Za to zdecydowanie użyteczną, choć niewidoczną stroną standardu Unicode są znaki „białe”: dzielenie wyrazów, niełaźliwy odstęp, niewidoczny łącznik, niewidoczny separator itd. (U+00AD, U+00A0, U+2060, U+2063). Bardzo przydatne są też, stanowiące część standardu Unicode, zasady tworzenia liter akcentowanych zaimplementowane w niektórych programach DTP.

Wydaje się jednak bezdyskusyjne, że w procesie przygotowywania publikacji standard Unicode jest obecnie powszechny i niezastąpiony. Tym bardziej, że zarówno nowe programy DTP, jak i nowo powstające fonty (w tym reedycje istniejących czcionek) domyślnie zakładają korzystanie ze standardu Unicode, a czasami wręcz wymuszają korzystanie z niego.

Unicode dla gisera

Proces ten zachodzi też w drugą stronę. Coraz szersze stosowanie standardu Unicode wymusza migracje fontów do formatów zgodnych z tym standardem. Jest to oczywiście pewne uproszczenie, ponieważ przyczyn migracji do nowych formatów, a także tworzenia nowych fontów w nowoczesnych formatach (mam tu na myśli głównie format OpenType) jest dużo więcej. Istnieją jednak głębokie powiązania pomiędzy możliwościami tego formatu a standardem Unicode – możliwość łatwego dostępu do dużej liczby znaków jest bardzo istotną cechą OpenType, a z drugiej strony ten format w pewnym sensie dopełnia i „poprawia” standard Unicode. To wzajemne powiązanie jest na tyle głębokie, że funkcjonuje nawet termin „font unikodowy”, oznaczający font OpenType (ew. TrueType) z szerokim repertuarem znaków, obejmującym co najmniej kilka systemów alfabetycznych.

Wprowadzenie nowych standardów zdecydowanie zmieniło zakres prac przy tworzeniu fontu. Oczywiście projektowanie kształtu czcionki jest nadal kluczowym elementem tego procesu. Wzrosło jednak znaczenie drugiego, programistycznego etapu pracy – można by wprowadzić tysiące znaków akcentowanych lub pochodnych (typu indeks górny lub dolny, wersja negatywowa), a także powiązań pomiędzy nimi tworzyć ręcznie, ale w takich zadaniach dużo lepiej sprawdzają się metody algorytmiczne.

Na pewno dużym „utrudnieniem” (pośrednio związanym z upowszechnieniem standardu Unicode) dla twórców fontów jest oczekiwanie użytkowników, by font zawierał wiele systemów alfabetycznych. Niewielu typografów ma wiedzę i doświadczenie potrzebne, by zaprojektować litery cyryliczne czy greckie, a jeszcze

mniej osób poradzi sobie jednocześnie z literami arabskimi czy znakami kanji. Obecnie wygląda na to, że tworzenie fontu jest raczej zadaniem dla międzynarodowego zespołu osób, niż dla jednego projektanta. Tę tendencję widać też na poziomie instytucjonalnym – wydawcami nowych fontów częściej są producenci oprogramowania (Microsoft, Apple, Adobe) niż firmy tradycyjnie tworzące fonty (Linotype, ITC, URW++).

Repertuar znaków

O ile przed rozpowszechnieniem standardu Unicode repertuar znaków był dobrze określony (zestaw liter z danego języka, ewentualnie trochę znaków akcentowanych i nieco symboli), a próby jego rozszerzenia nieliczne (fonty typu „ekspert”) i problemem było udostępnienie tych znaków w sposób wygodny dla użytkownika (specjalne indywidualne kodowania), o tyle teraz, w epoce Unicode, problemem staje się raczej rozsądne zawężenie zakresu znaków, które chcemy zawrzeć w foncie.

Praktycznie niemożliwe jest zrobienie fontu zawierającego 110 tys. znaków, a już zapewnienie ich spójności typograficznej jest chyba całkiem wykluczone. Najwięcej znaków ma GNU Unifont (63 tys.), chociaż trudno tu mówić o ich jakości. Solidniej zrobiony jest Arial Unicode MS, który ma ok. 40 tys. znaków. Dużo więcej fontów ma liczbę znaków rzędu dwóch tysięcy, co wydaje się rozsądniejsze – zwłaszcza jeśli dany font nie ma charakteru matematycznego. Na temat wymagań dotyczących fontów matematycznych – patrz niżej.

Dylemat, czy starać się zapewnić jak największe pokrycie standardu Unicode i jak najszerzy repertuar znaków czy raczej skupić się na jakości typograficznej i spójności wizualnej znaków, jest obecnie bardzo podstawową decyzją projektową przy rozpoczynaniu prac nad fontem.

O ile kody z pierwszych 7 zakresów standardu (U+0020 do U+036F) na pewno warto zawrzeć w foncie, o tyle dalsze zakresy (i systemy alfabetyczne) zdecydowanie poszerzają zakres prac, tym bardziej że złożoność zadania wzrasta proporcjonalnie (co najmniej) do kwadratu liczby znaków ze względu na zależności pomiędzy znakami. Na przykład w obecnych technologiach pełen kerning pomiędzy wszystkimi znakami z różnych języków jest bardzo trudny do realizacji i nieefektywny.

W standardzie Unicode jest wiele takich zakresów, które nadają się raczej do bardzo specjalistycznych fontów i chyba nie trzeba ich umieszczać w czcionkach powszechnego użytku, na przykład: U+4D??, czyli „heksagramy Yijing”, U+1F0??, czyli „kostki domina” czy przykłady przedstawione na rysunku 2. Do tego docho- dzą zakresy typu U+277? (*DINGBAT CIRCLED SANS-SERIF DIGIT* – czyli: „symbole bezszeryfowych cyfr w okręgach”), które są świadectwem zaszłości zastygłych w standardzie Unicode.

Unicode a typografia specjalistyczna

Osobnym obszarem standardu Unicode są jego zastosowania specjalistyczne. Takim bardzo ważnym i dynamicznie rozwijającym się zastosowaniem Unicode jest matematyka, a dokładniej zapis wzorów matematycznych (które stanowią część matematyki, ale także fizyki, chemii i wielu innych dziedzin nauki).

Zdecydowanie jednak odradzałbym dążenie do zapewnienia jak największego pokrycia zakresów standardu Unicode kosztem jakości typograficznej. W razie potrzeby zawsze można wstawić „jakiś” symbol z innego fontu, a estetyka składu matematyki zależy głównie od najpopularniejszych symboli, a nie od tych rzadkich.

Warto w tym miejscu zwrócić uwagę, że istnieją też duże obszary działalności człowieka, które standard Unicode pomija. O ile matematyka została szeroko uwzględniona w standardzie, to tak obszerna dziedzina sztuki jak muzyka jest potraktowana bardzo skromnie – tylko jeden zakres U+1D1??, dodany zresztą dopiero w wersji 3.1. standardu Unicode. W dodatku te znaki nie wystarczą do zapisania nawet prostej melodii – nie pozwalają przekazać informacji o wysokości nuty czy dokładnego czasu trwania dźwięku.

Innym przykładem istotnej niekompletności standardu Unicode jest język migowy – dodanie znaków do zapisu tego języka znajduje się dopiero w sferze planów, nawet nie projektowej. Trochę dziwny jest także status znaków Braille’a: w standardzie Unicode uwzględniono wszystkie 64 „układy punktów”, a nie litery, symbole – wbrew zapewnieniom o uwzględnianiu semantyki i nie zajmowaniu się reprezentacją graficzną.

Wady standardu Unicode

Oprócz wad (czy mówiąc delikatnie: nie zawsze pożądanymi właściwościami) wymienionych w poprzednich punktach, chciałbym też odnieść się do stanu realizacji poszczególnych założeń standardu:

- *uniwersalny repertuar* – mimo rozległości obsługiwanych znaków cały czas pozostaje niedosyt związany z brakiem kompletności zakresów, na przykład symbole literowe w matematyce – niektóre odmiany symboli zawierają grecki alfabet i cyfry, a inne nie zawierają albo jednego, albo drugiego;
- *porządek logiczny* – ten postulat został spełniony w najmniejszym stopniu ze względu na przyrostowość specyfikacji – na przykład matematyczne symbole pisma tablicowego (*double struck letters*) znajdują się w dwóch zakresach: litery H, I, R, Z, C w jednym, a reszta w zupełnie innym (por. rysunek 3); tak samo strzałki, które są podzielone na cztery zakresy (a do tego tak, że strzałka potrójna w prawo jest w innym rejonie niż taka sama strzałka w lewo);
- *efektywność* – brak logicznego porządku powoduje komplikacje także w tym zakresie: nie da się prosto i algorytmicznie określić, które znaki powinny być utożsamiane dla potrzeb wyszukiwania, czy jaki jest porządek alfabetyczny znaków; te informacje są oczywiście dostępne, ale tylko w postaci specjalnych tablic;
- *ujednoliczenie* – niestety, tu też widać słabości standardu Unicode, wiele znaków jest powtórzonych, i to nie tylko pojedyncze, jak: U+00B5 – *symbol mikro* i U+03BC – *litera mu*, ale też całe bloki, jak U+FF?? – *Halfwidth and Fullwidth Forms* będący kopią liter i cyfr łańciskich dla użytku wschodnioazjatyckiego;
- *znaki, a nie kształty* – sytuacja daleka od założeń: znaki w indeksie górnym (U+207?), zestawienia liter – ligatury, liczby rzymskie (U+216?) czy wspomniane już cyfry w kółeczkach (U+278?);

- *znaczenie (semantyka)* – z dwóch powyżej omówionych punktów wynika, że także ten postulat jest daleki od realizacji; konsorcjum Unicode chętnie powołuje się na to założenie, uzasadniając odmowę dodania jakiejś kategorii znaków, ale jednocześnie unika go przy analizie sensowności decyzji podjętych w przeszłości;
- *stabilność* – w tym zakresie sytuacja znacząco się poprawiła, ale na przykład ślady zamieszania związanego z pomieszaniem akcentów *cedilla* i *przecinek* pod literami *T, t, S i s*, co zostało poprawione dopiero w wersji 3.0 standardu, wciąż jeszcze trwają;
- *przenośność* – tu zastrzeżenia mogą dotyczyć tylko źle przygotowanych tekstów, na przykład takich, jakie zostały wspomniane w punkcie „Unicode dla zecera”.

Propozycje rozwiązań alternatywnych

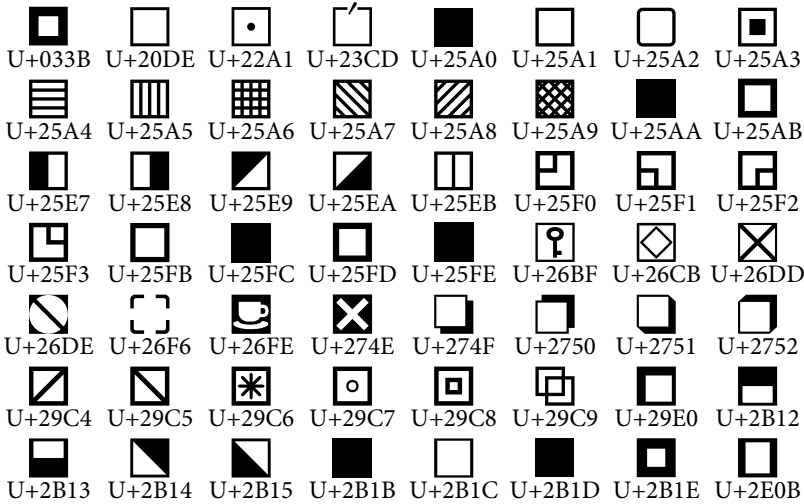
Jak widać z powyższych rozważań, standard Unicode ma kilka słabych punktów – niestety, większości z nich nie da się usunąć w ramach tego standardu, ponieważ wynikają one z podstawowego założenia „ponumerowania wszystkich znaków” oraz z przyjętej od początku polityki „zgodności wstecz”, sięgającej aż do standardu ASCII.

Niewątpliwie należy odciąć się od zasady „numerowania” jako podstawy służącej do przekazywania szeroko rozumianej informacji tekstowej. Standard Unicode jest nazywany ostatecznym kodowaniem – i rzeczywiście, tworzenie lepszego standardu kodowania nie jest celowe. Warto myśleć o innych sposobach zapisywania dokumentów.

Wartą rozważenia propozycją byłoby oparcie standardu wymiany dokumentów (już nie „kodowania”) na nazwach znaków. Standard Unicode określa także nazwy znaków i zawiera wskazówki co do ich tworzenia. Niestety, nazwy nie zawsze były tworzone konsekwentnie, co doprowadziło do tego, że np. pod nazwą nie zawsze jest to, czego byśmy się spodziewali – zob. rysunek 4. Gdyby jednak po prostu określić protokół (zasady) tworzenia unikalnych, jednoznacznych i łatwych do interpretacji nazw znaków, najlepiej z wykorzystaniem hierarchicznych przestrzeni nazw, to można by stworzyć standard zastępujący lub może uzupełniający Unicode.

Można to sobie wyobrazić jako coś podobnego do systemu nazw (DNS) w Internecie – z tym że nazwa byłaby tłumaczona nie na numer IP, a na konkretny znak (albo na algorytm, jak go skonstruować, np. w przypadku mało popularnych znaków akcentowanych, czy symboli geometrycznych). Nazwa mogłaby też zawierać opcjonalne elementy uściślające kontekst wykorzystania znaku – np. język, w którym dany znak jest użyty, czy wręcz wygląd tego znaku (font, którym jest złożony czy wariant stylistyczny).

Drugim, dalej idącym rozwiązaniem byłoby przyjęcie, że znak nie jest *numerem* ani nawet *nazwą*, ale raczej *obiektem*, zawierającym dużo więcej informacji o tym znaku. W standardzie Unicode występuje załączek tej koncepcji w postaci bazy UCD (*Unicode Character Database* [4]). Gdyby taki zbiór obiektów „znakowych” stał się centrum nowego standardu, a numer (być może niekoniecznie jeden – można sobie wyobrazić kilka kodowań do różnych zastosowań) i nazwa służyły jako



Rys. 4. Symbole Unicode'owe o nazwach zawierających słowo *square* – czyli kwadrat. W tym 4 białe i 6 czarnych kwadratów. A bardzo podobne są także różne formy: *rectangle*, *bar* czy *shade*

klucz dostępu (indeks według nomenklatury bazodanowej), mielibyśmy do czynienia z nową jakością, zarówno w zakresie obiegu dokumentów, jak i przygotowywania publikacji do druku (w tym typografii).

Trzecim możliwym kierunkiem zmian jest uproszczenie struktury dokumentu wejściowego (i co za tym idzie, ograniczenie środków potrzebnych do jego zapisania) oraz przeniesienie części informacji (zwłaszcza typograficznej) do programu fontowego. Tak naprawdę niektóre decyzje należałoby podejmować dopiero na etapie łączenia tekstu z fontem (np. dołączanie akcentów do liter, wybór początkowego lub końcowego wariantu litery, ligatury, indeksy górne i dolne). Obecnie są one zawarte w standardzie Unicode (choć nie w całości). Najlepiej byłoby, gdyby takie decyzje dało się wyrazić w sposób programowalny (definiowany przez autora, typografa czy użytkownika). Dopiero możliwość użycia algorytmu do wyrażenia pewnych zależności czy decyzji pozwala w pełni wykorzystać zalety komputerowego przetwarzania tekstów.

Oczywiście w powyższych propozycjach nadal jest miejsce na jakies „kodowanie znaków” – tylko że powinno ono być uzupełnieniem, używanym przy prostych zadaniach, a nie naczelną zasadą. W takiej roli standard Unicode byłby nadal użyteczny (choć należałoby się zastanowić nad powrotem do jego wersji 16-bitowej).

Podsumowanie

Na pewno standard Unicode jest dużym krokiem naprzód w stosunku do rozwiązań poprzedzających, ale jak wynika z powyższych rozważań, daleko mu do rozwiązania „ostatecznego”, nie należy spodziewać się, że jest długotrwałe i rozwojowe. Wręcz odwrotnie, wydaje się, że jest wycinkowe i że warto szukać rozwiązań alternatywnych. Ze względu jednak na rozmiar zadania prace nad nowymi standardami wymagają zaangażowania sporego zespołu. Przedstawione propozycje nie są goto-

wymi rozwiązaniami, ale należy mieć nadzieję, że w połączeniu z wskazaniem wad obecnego rozwiązania mogą być one użyteczne w dalszych pracach nad fundamentalnymi założeniami wymiany informacji tekstowej, uwzględniającymi możliwości rozwijającej się dynamicznie technologii.

Cytowana literatura

1. The Unicode Consortium, *Who we are*.
<http://www.unicode.org/consortium/consort.html>, dostęp 17.04.2013.
2. "Unicode.": *the Universal Telegraphic Phrase-Book. A Code of Cypher Words for Commercial, Domestic, and Familiar Phrases in Ordinary Use in Inland and Foreign Telegrams. With a List of Prominent Commercial Firms who are Unicode Users*. Cassell & Company Limited 1889. <http://books.google.pl/books?id=W2U0AAAAMAAJ>, dostęp 17.04.2013.
3. Becker Joseph D., *Unicode 88*. Palo Alto: Xerox Corporation 1988.
<http://www.unicode.org/history/unicode88.pdf>, dostęp 17.04.2013.
4. The Unicode Consortium, *Unicode Character Database 6.2.0*. The Unicode Consortium 2012. <http://www.unicode.org/Public/UNIDATA/UCD.zip>, dostęp 17.04.2013.

Abstract

Unicode standard in typography

The Unicode standard is designed as a framework for the worldwide interchange and processing of documents written in various languages and from various technical disciplines. The Unicode standard, however, does not define glyph images and neglects typography. Nevertheless, it has been adapted as the *de facto* standard in typography. The advantages of such a universal standard are obvious, but its objectives result in serious drawbacks in typographical applications. Therefore our conclusion is that it will be advisable to explore alternative solutions.