

# PLAIN

## Opis pakietu M<sub>E</sub>X

Bogusław Lichoński

Z wielką przyjemnością przedstawiam pierwszy z prawdziwego zdarzenia dwujęzyczny format dla T<sub>E</sub>X-a o leśnej nazwie M<sub>E</sub>Xw którym obok języka angielskiego pojawia się język polski.

M<sub>E</sub>X 1.03

Już 18 grudnia 1992 roku autorzy pakietu — panowie Bogusław Jackowski i Marek Ryćko — udostępnili nam wszystkim wersję 1.03 pakietu M<sub>E</sub>X. Składa się on z fontów w postaci źródłowych plików METAFONT-owych, zawierających polskie znaki diakrytyczne i używane w języku polskim znaki cudzysłowów, formatów w wersji źródłowej, zwanych M<sub>E</sub>X i I<sub>A</sub>M<sub>E</sub>X (odpowiedników formatów plain i lplain) oraz polskich wzorców dla T<sub>E</sub>X-owego algorytmu dzielenia wyrazów. Całość jest oprogramowaniem typu *public domain*.

Oto szczegółowa lista plików pakietu M<sub>E</sub>X:

- zbiór plików formatu M<sub>E</sub>X; mex.tex, mex1.tex, mex2.tex, mexconf.tex, polhyph.tex;
- format I<sub>A</sub>M<sub>E</sub>X; lamex.tex;
- zbiór plików wsadowych i pomocniczych do generowania obu formatów w środowisku T<sub>E</sub>X-a 3.x z implementacji emT<sub>E</sub>X-a. gm-lamex.bat, gm-mex.bat, gl-lamex.bat, gl-mex.bat, maz2pl.tcp, maz2pl.tpc, lat2pl.tcp, lat2pl.tpc.

Nad pakietem pracowali także: pani Hanna Kołodziejska, autorka polskich wzorców dzielenia wyrazów oraz pan Roman Tomaszewski, wybitny polski typograf, prezes polskiego oddziału A-Typ-I (*Association Typographique Internationale*).

zalecana jest notacja bezprefiksowa, bieżącym językiem jest język polski, 256 znaków w foncie, frenchspacing

Z chwilą rozpoczęcia pracy z formatem M<sub>E</sub>X obowiązują nieco odmienne zasady niż z formatem plain. Po pierwsze bieżącym językiem jest język polski. Po drugie obowiązującym sposobem notacji polskich znaków w plikach źródłowych jest

notacja bezprefiksowa (bezciachowa) i może to być standard Mazovia albo Latin 2. Po trzecie stosowanym układem polskich znaków diakrytycznych w fontach jest układ PL. Nowością jest wykorzystanie możliwości T<sub>E</sub>X-a 3.x i wprowadzenie układu 256 znaków w foncie. Po czwarte obowiązującym sposobem spacjowania po znakach przystankowych jest frenchspacing, czyli wielkość odstępów między wyrazami nie zależy od tego, jaki znak poprzedza odstęp.

Użytkownik może zmienić powyższe zasady modyfikując plik konfiguracyjny mexconf.tex, a następnie ponownie wygenerować format.

mamy możliwość tworzenia polskich komend —  
`\Ściąga, \wskazówka`

Mój pierwszy kontakt z M<sub>E</sub>X-em był dla mnie bardzo miły. Denerwująca mnie od początku używania T<sub>E</sub>X-a notacja ciachowa przy pisaniu polskich liter została zastąpiona przez wygodne umieszczanie w tekście polskich znaków diakrytycznych zapisywanych jako zwykłe znaki o kodach w preferowanej przeze mnie Mazovii. Pozwala to na bardzo wygodne składanie tekstów! Polskie litery widoczne są na ekranie w trakcie przygotowywania tekstu, ponadto możliwe jest definiowanie polskich komend T<sub>E</sub>X-owych, jak choćby `\Ściąga, \wskazówka`.

`\prefixing id/x ku s/lo/ncu`  
`\nonprefixing idź ku słońcu`

Aby zachować zgodność z L<sub>E</sub>X-em (wcześniejszą adaptacją T<sub>E</sub>X-a, również autorstwa panów BJ & MR) M<sub>E</sub>X posiada komendę `\prefixing`, która umożliwi pisanie tekstów w notacji prefiksowej (notację bezprefiksową przywraca `\nonprefixing`).

Ponieważ znak `'/` w notacji prefiksowej wykorzystany jest do oznaczania polskich liter, konieczne było wprowadzenie konwencji notacyjnej pozwalającej uzyskać w składzie symbol „ciach”. W M<sub>E</sub>X-u przyjęto naturalną zasadę uzyskania znaku „ciach” przez `'/'`. (Można także używać dłuższej komendy `\normalslash`.) Należy zwrócić uwagę, że w czasie pisania do pliku za pomocą komendy `\write` makra są rozwijane. Dotyczy to w także makra `/` i trzeba na to zwracać uwagę przy pisaniu do plików, które mają być potem czytane przez T<sub>E</sub>X-a. Użyteczne też mogą być komendy `\emulateplain` i `\emulateLaTeX` służące do emulacji formatów plain i L<sup>A</sup>T<sub>E</sub>X odpowiednio wewnątrz formatów M<sub>E</sub>X i I<sub>A</sub>M<sub>E</sub>X.

```
\language\english
I wanna go home
\language\polish
Ja chcę iść do domu
```

Zaletą M<sub>TeX</sub>-a jest możliwość przełączania między sposobami dzielenia wyrazów. W praktyce dosyć często w dokumentach pojawiają się fragmenty choćby w języku angielskim. „Uruchomienie” reguł dzielenia wyrazów angielskich dokonujemy komendą `\language\english`. Przełączenie na reguły polskie dokonujemy poleceniem `\language\polish`.

```
I like "Times"
Ja lubię „Wprost”
Je préfère «Le Monde»
```

Autorzy M<sub>TeX</sub>-a nie zapomnieli o polskich cudzysłowach. Bardzo częstym błędem jest używanie w polskich tekstach anglosaskich cudzysłowów. Polski cudzysłów otwierający uzyskać można przez napisanie dwóch znaków `„`, w składzie uzyskamy `„`, podobnie jest z cudzysłowem zamykającym, to znaczy pisząc `”` (lub `»`) otrzymamy `”`. Niekiedy w języku polskim używany jest też inny rodzaj cudzysłowów, tak zwane cudzysłowy «francuskie». Lewy francuski cudzysłów oznaczany jest dwoma znakami mniejszości, a prawy dwoma znakami większości.

Przy kompilacji znaków `„`, `”`, `<<` i `>>` na odpowiadające im w składzie cudzysłowy wykorzystywany jest T<sub>EX</sub>-owy mechanizm ligatur. Oczywiście informacje o ligaturach zostały zawarte w METAFONT-owych źródłach fontów PL.

```
\plqq Pani\prqq      „Pani”
\flqq Elle\frqq      «Elle»
```

Mechanizmu wstawiania cudzysłowów za pomocą ligatur nie zastosowano w fontach stosowanych do składu formuł matematycznych oraz fontach nieproporcjonalnych (np. włączane komendą `\tt`). Cudzysłowy można wtedy uzyskać przez użycie komend:

- `\plqq`, `\prqq` — cudzysłowy polskie,
- `\flqq`, `\frqq` — cudzysłowy francuskie.

```
\layoutpl
\layoutlatintwo
\layoutpone
\layoutmazovia
```

Być może niektórym użytkownikom T<sub>EX</sub>-a zdarzyło się wykorzystywać w jednym dokumencie fonty różnego pochodzenia, ale czy udawało się robić taki koktajl z fontami o różnych układach polskich znaków? M<sub>TeX</sub> posiada taką możliwość. Możliwe jest to dzięki komendzie `\layout{xxx}`, gdzie `xxx` ∈ {`pl`, `mazovia`, `latintwo`, `pone`}.

Przedstawiony wyżej zbiór posiada cztery elementy, z których jedynie `pone` (skrót od `polishone` czyli P1) może być niezrozumiały. Użycie komendy `\layout{pone}` powoduje przygotowanie M<sub>TeX</sub>-a do pracy z fontami P1, poprzednikami fontów PL. Jak widać M<sub>TeX</sub> przygotowany jest do pracy z fontami w układzie PL jako podstawowymi oraz z fontami zawierającymi polskie litery zgodne z kodem Mazovia, Latin 2 oraz P1. Należy zwrócić uwagę, że komenda `\layout` nie „włącza” samego fontu, jedynie przygotowuje system do pracy z fontami w odpowiednim układzie.

W M<sub>TeX</sub>-u polskie znaki mają kategorię „litera”

Przyjrzyjmy się dokładnie komendzie `\layout`. Wykonanie jej polega na przypisaniu polskiemu znakom diakrytycznym kategorii „litera”. Jeżeli używamy układu innego niż PL, polskiemu znakom przypisana zostaje kategoria „aktywny” i traktowane są one jako makra rozwijające się do odpowiednich liter. Oprócz tego dokonuje się odpowiednia zamiana tablicy `\sfcode` oraz utworzenie tablic `\uccode` i `\lccode` zgodnych z układem fontu. W notacji prefiksowej odpowiednio zmieniają się też znaczenia komend `\plqq`, `\prqq`, `\flqq`, `\frqq`. Pozycjom polskich znaków zostają przypisane odpowiednie kody matematyczne. Ważnym efektem jest włączenie reguł przenoszenia wyrazów w języku polskim (zgodnych z układem fontu), jeśli nie ma odpowiednich reguł, włączany jest zakaz dzielenia wyrazów. Zmienia się także znaczenie symbolu `\polish`.

Niezależnie od obowiązującego układu fontu i znaczenia symbolu `\polish`, symbole:

- `\pllanguage`,
- `\mazovialanguage`,
- `\latintwolanguage`,
- `\ponelanguage`

oznaczają numery języków w sensie T<sub>EX</sub>-a, czyli numery odpowiednich reguł dzielenia wyrazów.

Oczywiście T<sub>EX</sub> „widzi” zasady dzielenia wyrazów zależne zarówno od języka, jak i od układu

liter w bieżącym foncie. Jeżeli więc użytkownik „włącza” font o innym układzie (innych kodach) polskich znaków diakrytycznych, to towarzyszyć temu powinno przełączenie T<sub>E</sub>X-a na inne reguły przenoszenia wyrazów.

Podczas generowania formatu `mex.fmt` domyślnie wbudowywane są reguły przenoszenia wyrazów w języku angielskim i polskim.

Użytkownik — zależnie od wielkości dostępnej pamięci w komputerze — może przystosować format do korzystania z reguł przenoszenia dla innych układów polskich fontów. Wystarczy dokonać prostej modyfikacji pliku `mexconf.tex`.

<code>żółto\=niebieski</code>	<code>żółto-</code> <code>-niebieski</code>
-------------------------------	--

Polskie reguły ortograficzne są czasami sprzeczne z anglosaskim. Na przykład zaleca się by wyrazy łączone, takie jak „żółto-niebieski”, przenosić w miejscu łączenia. W języku angielskim nie powtarza się łącznika na początku drugiego wiersza. Ale w polskim dywiz należy powtórzyć na początku nowej linii. Dlatego w polskiej wersji została wprowadzona komenda `\=`, która użyta w miejsce łącznika (żółto-niebieski) powoduje dzielenie zgodne z opisaną regułą. Ponieważ jednak w formacie `plain` komenda `\=` oznacza akcent „macron” dlatego format M<sub>E</sub>X zastępuje tę niedogodność makrem `\macron`.

$x_{\text{średnia arytmetyczna}} \geq x_{\text{średnia geometryczna}}$ $a \frac{x}{\ln x} \leq \pi(x) \leq b \frac{x}{\ln x}$
---

M<sub>E</sub>X dba również o matematykę. Co prawda modyfikacje w stosunku do formatu `plain` są tu niewielkie, jednak interesujące. Nowością jest używanie polskich liter (w dowolnej notacji) w formułach matematycznych w indeksach górnych i dolnych. Stosowane w Polsce znaki relacji mniejsze-równe ( $\leq$ ) i większe-równe ( $\geq$ ) różnią się od znaków zaprojektowanych przez D. E. Knutha. Fonty PL (ściślej PLSY) zawierają polskie znaki relacji. Uzyskujemy je za pomocą komend: `\xleq` lub `\xle` oraz `\xgeq` lub `\xge`.

<p>„polskie” funkcje: <code>\tg</code>, <code>\ctg</code>, <code>\tgh</code>, <code>\ctgh</code> oraz <code>\arc\sin</code>, <code>\arc\cos</code>.</p>
---

Zauważmy, że w Polsce używa się do nazwania funkcji trygonometrycznych symboli „tg”, „ctg”, „tgh” i „ctgh” a nie „tan”, „cot”, „tanh” i „coth”. Twórcy M<sub>E</sub>X pomyśleli i o tym tworząc komendy `\tg`, `\ctg`, `\tgh`, oraz `\ctgh`. Jest to niby tylko zmiana nazwy, jakże jednak kojarząca się nam ze swojskim zapisem. Dodatkowo wprowadzona została komenda `\arc`, która w połączeniu z powyższymi (`\arc\sin`, `\arc\cos`, itd.) daje w składzie `arc sin`, `arc cos` itd. Jak widać między słowem „arc” i nazwą funkcji istnieje odpowiedni odstęp.

Na zakończenie warto wspomnieć, że obok formatu M<sub>E</sub>X pakiet zawiera także format L<sub>A</sub>M<sub>E</sub>X, który jest polskim odpowiednikiem angielskiego formatu `lplain`. Autorzy M<sub>E</sub>X-a zadbali więc o wszystkich zwolenników L<sup>A</sup>T<sub>E</sub>X-a.

Podsumowując, M<sub>E</sub>X okazał się bardzo wygodnym narzędziem do składania polskich tekstów. Nie jest oczywiście ideą nową, tego typu formaty już dawno powstały w innych krajach. M<sub>E</sub>X zbliżył nas więc nieco do Europy, dając komfort pracy. Aby nie tworzyć pustych zdań powiem, że M<sub>E</sub>X bardzo podoba mi się, używam go od lutego 1992 roku w środowisku implementacji emT<sub>E</sub>X. Takie połączenie wydaje mi się bardzo dobre i mogę je wszystkim polecić.

Formaty M<sub>E</sub>X i L<sub>A</sub>M<sub>E</sub>X wygenerowałem przy użyciu kompilatora `tex286.exe` wersja 3.0 [3a], zaś źródła fontów PL skompilowałem METAFONT-em `mf286.exe` wersja 2.0 [3a], przy użyciu programu `mfjob.exe` wersja 1.1f. Wszystkie powyższe programy pochodzą z pakietu emT<sub>E</sub>X-a E. Mattes’a.