ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria II: WIADOMOŚCI MATEMATYCZNE XXVII (1986)

Ministerstwo Nauki i Szkolnictwa Wyższego

MAREK J. LAO (Warszawa)

Sztuka TEXowania

Składanie tekstów matematycznych jest pracą bardzo uciążliwą. Zecerzy narzekają na olbrzymią pracochłonność przygotowywania do druku książek, artykułów czy innych publikacji. Matematycy są czasami zmuszani do pisania $x^{1/2}$ zamiast $\sqrt{x}$. I najczęściej teksty prac opublikowanych w czasopismach są zupełnie różne od tego, co wymarzył sobie autor. Zresztą każdy, kto kiedykolwiek przepisywał na zwykłej maszynie do pisania swój własny artykuł, wie, ile ciężkiej pracy i pomyłek jest spowodowanych przez konieczność ręcznego dopisywania indeksów, specjalnych symboli, liter alfabetu greckiego, gotyckich czy pisanych. Wydaje się, że te problemy są bliskie końca. Chyba już wkrótce w wielu krajach komputerowe składanie tekstów matematycznych zastąpi pracę zecerów i — jeśli tylko komputery osobiste staną się tak rozpowszechnione jak maszyny do pisania — również skończy się ręczne dopisywanie symboli i kolorowe podkreślanie tekstu dla zaznaczenia informacji przeznaczonych dla drukarni.

Pierwsze teksty składane przy użyciu komputera zaczęły się pojawiać w latach sześćdziesiątych. Było to możliwe dzięki wprowadzeniu różnych maszyn do fotoskładu. American Mathematical Society oraz National Science Foundation rozpoczęły wówczas badania dotyczące efektywnego i estetycznego składania tekstów. Do badań dołączyły wkrótce Digital Equipment Corporation oraz Xerox Palo Alto Research Laboratories. Zaczęto produkcję coraz doskonalszych urządzeń drukarskich opartych na zasadzie fotoskładu rastrowego.

Fotoskład rastrowy jest podobny do odtwarzania obrazów przesyłanych na odległość, na przykład obrazów uzyskanych ze statków kosmicznych. Płaszczyzna przeznaczona do zadrukowania podzielona jest na mikroskopijne kwadraty. Maszyna do fotoskładu dostaje informację, które z tych kwadratów rastrów mają być czarne (zadrukowane), a które mają pozostać białe. Na podstawie tej informacji, maszyna naświetla przy użyciu wiązki elektronowej lub laserowej matrycę wykonaną z materiału fotoczułego. Taka matryca może być teraz użyta do drukowania

książek i czasopism. Fizyczne własności farby drukarskiej powodują, że niewielkie ostrości obrazu z matrycy są „wygładzane". Tak więc ostateczny rezultat całego procesu składa się z liter i symboli, których kontury są gładkie. Na podobnej zasadzie opierają się również komputerowe drukarki mozaikowe; te jednak zawierają za mało „punktów" rastrowych przeznaczonych na poszczególne litery, byśmy mogli mówić o wysokiej jakości tekstów drukowanych za ich pomocą (1).

Jak więc widzimy, fotoskład rastrowy oparty jest na przygotowaniu odpowiednich danych — olbrzymiej macierzy kwadratowej zawierającej informację o tym, czy dany „punkt" rastra ma być biały czy czarny, tzn. tablica ta zawiera informacje binarne — „tak"/„nie" lub w żargonie komputerowym 1/0. Ten fakt oraz naprawdę wysoka jakość druków otrzymanych w ten sposób sprawiły, że tematyką komputerowego składania tekstów zajął się jeden z najwybitniejszych matematyków-informatyków światowych, Donald E. Knuth, profesor Stanford University, specjalista w zakresie matematyki kombinatorycznej, autor podstawowej monografii w dziedzinie analizy algorytmów *The art of computer programming*.

Knuth rozpoczął swoje prace w zakresie komputerowego składania tekstów matematycznych w 1977 roku. Opracował dwa systemy programów: METAFONT oraz TEX. METAFONT służy do definiowania alfabetu oraz symboli. Za jego pomocą można zbudować odpowiednie fragmenty macierzy zawierającej dane dla fotoskładu. Oczywiście mało który autor artykułu interesuje się tą częścią systemu; tym bardziej, że system dostarcza tak dużej liczby różnych krojów pisma i znaków, że rzadko kiedy pojawia się prawdziwa potrzeba zdefiniowania dodatkowych symboli(2). I tak na przykład dostępny jest alfabet grecki do użycia w formułach matematycznych, druk pochyły (kursywa), półgruby, tak zwana czcionka bezszeryfowa oraz kapitaliki. Oprócz tego możemy używać liter gotyckich, pisanych oraz symboli specjalnych. W razie potrzeby możemy wydrukować tekst w taki sam sposób jak na maszynie do pisania, gdzie każda litera ma taką samą szerokość.

Częścią, która interesuje większość użytkowników, jest TEX. Nazwa pochodzi od greckiego słowa $\tau\varepsilon\chi$ oznaczającego sztukę i technikę. Tak więc zgodnie z sugestiami autora, TEX powinno być wymawiane jako „tech" (jednakże większość osób, z którymi rozmawiałem, wymawia to słowo jako „tek"; używa się go też jako czasownika na określenie składania tekstu przy użyciu tego systemu). TEX, a w szczególności jego wersja tworzona przez Michaela Spivaka na zlecenie American Mathematical

(1) Na tej zasadzie oparta jest technika monophoto, którą wydrukowano kilka zeszytów *Wiadomości Matematycznych* w latach 1981–1983 (*przyp. Red.*).

(2) W czasie pobytu w Indiana University, Bloomington zdefiniowałem swoje własne symbole odpowiadające literom alfabetu dla języka polskiego jedynie za pomocą składania symboli standardowych.

Society — tzw. AMS TEX, jest bardzo przystępny i w zasadzie nie wymaga od użytkownika żadnej wiedzy informatycznej. Oczywiście odrobina kultury informatycznej jest przydatna. Dobrze by było, gdyby użytkownik umiał korzystać z terminala, zakładać i poprawiać pliki zawierające teksty. Tych umiejętności można się nauczyć w ciągu godziny! Dlatego też zachęcam wszystkich, którzy mają możliwość próby użycia TEXa, do zrobienia tego. System jest dość popularny na uniwersytetach amerykańskich, a pisanie artykułów za jego pomocą naprawdę sprawia ogromną satysfakcję. By uzasadnić tę opinię, omówię kilka przykładów użycia TEXa do tekstów matematycznych.

Najprostszym przykładem jest napisanie formuły matematycznej, np. $y = 3(x-1{,}5)/(-2|x|+1)$, wewnątrz akapitu. Tekst samego akapitu może być pisany w zasadzie dowolnie. TEX usunie nadmierne odstępy pomiędzy wyrazami, podzieli tekst na wiersze oraz wprowadzi odpowiednie odległości tak, by szerokości kolejnych wierszy były jednakowe oraz by tekst ostatniego wiersza akapitu był przesunięty do lewego marginesu. Oczywiście TEX zadba o to, by odległość pomiędzy kolejnymi wierszami umożliwiała wydrukowanie tekstu bez nakładania się symboli napisanych na różnych poziomach, oraz o to, by pierwszy wiersz akapitu był wcięty.

Podział akapitu na wiersze jest automatyczny; TEX będzie próbował zrobić to tak, by tekst wyglądał estetycznie oraz by nie trzeba było przenosić części wyrazów z wiersza do wiersza. Czasami przeniesienie jest konieczne; wówczas TEX zastosuje reguły przenoszenia pisowni (angielskiej) i odpowiednia część zostanie przeniesiona do następnego wiersza. Można poinformować system o tym, jak podzielić trudne wyrazy (w szczególności zastosować reguły pisowni polskiej), i wtedy, jeśli przeniesienie jest konieczne, TEX przedzieli wyraz w ten sposób we wskazanym miejscu. Zauważmy, że dzięki komputerowemu przetwarzaniu tekstu przed drukowaniem, treść całego akapitu wpływa na to, jak zostanie on podzielony na wiersze. Akapit kończy się specjalnym poleceniem \par (skrót dla *paragraph*). Tekst występujący dalej znajdzie się w nowym akapicie.

Wróćmy jednak do naszego zadania. Załóżmy, że chcemy wprowadzić formułę postaci $y = 3(x-1{,}5)/(-2|x|+1)$ do tekstu akapitu takiego jak ten. Jest to możliwe przez przełączenie systemu do trybu matematycznego. Wszystko, co jest ujęte w dwa kolejne znaki specjalne \$, jest traktowane jako formuła matematyczna. TEX wydrukuje wszystkie litery za pomocą druku pochyłego, tak więc „wyrazy" odpowiadają iloczynowi zmiennych. Zastosowane będą specjalne reguły umieszczania odstępów pomiędzy różnymi częściami formuły. Te zasady są bardzo naturalne, nie ma więc potrzeby informowania systemu, jak to zrobić. W naszym przykładzie wystarczy napisać:

... postaci \$$y = 3(x-1{,}5)/(-2|x|+1)$\$ do tekstu ...

Wewnątrz formuł matematycznych możemy również korzystać z liter alfabetu greckiego, symboli specjalnych i wielu innych. Dla przykładu podamy jeszcze zapis formuły $x^2 = \sqrt{\cos a}$. Może ona być uzyskana następująco:

... podamy jeszcze $x\uparrow 2 = \sqrt{\cos\alpha}$. Ta ...

Należy jednak sprawdzić, czy symbol \cos istnieje w wersji systemu, z której korzystamy. Nawiasy { } służą TEXowi jako informacja o tym, jak należy grupować (rozumieć) dany zapis. I tak na przykład formuły ${x\uparrow i}\uparrow j$ oraz $y\uparrow{k\uparrow l}$ pojawiają się jako x^{ij} oraz y^{k^l}. W szczególności $e\uparrow{-\cos\alpha}$ da nam: $e^{-\cos\alpha}$.

TEX stara się wydrukować formułę matematyczną w całości w jednym wierszu. Jeśli nie jest to możliwe, będzie próbował przenieść część formuły za znakami relacji =, <, a w następnej kolejności za operacjami +, − itd. Tak więc duża część naszej pracy odpada. Jeśli wiersze formuły są skomplikowane (np. zawierają dużo wykładników), to system dodaje odpowiednio większy odstęp dla oddzielenia tych wierszy od siebie.

Nieco inna wersja drukowania formuł matematycznych umożliwia wyodrębnienie formuły ze zwartego tekstu akapitu. Ujęcie formuły pomiędzy podwójne znaki $$ oznacza, że formułę należy wystawić do oddzielnego wiersza (*display*). Jest ona wówczas drukowana poza właściwym tekstem akapitu, otoczona większym odstępem, a także — jeśli zawiera indeksy, wykładniki lub symbole specjalne (np. znaki całkowania) — są one drukowane przy użyciu większych czcionek niż te, które są stosowane dla formuł wplecionych w tekst. Formuła jest więc czytelniejsza. Jeśli więc chcemy wydrukować w oddzielnym wierszu

$$\Phi = 2^{2^{2^2}},$$

to nasz zapis powinien mieć następującą postać:

... w oddzielnym wierszu $$\Phi = 2\uparrow{2\uparrow{2\uparrow{2\uparrow 2}}}$$ to ...

W formułach matematycznych wystawionych do oddzielnego wiersza możliwe jest również drukowanie wzorów bardzo złożonych tak, by określone fragmenty były wyrównane, drukowane bezpośrednio pod sobą. Dla przykładu rozpatrzmy:

$$\$\backslash\text{align } x+x+2-2x \;\&\; = 2x+2-2x\backslash\text{cr} \;\&\; = 2\backslash\text{cr}\$\$.$$

Znak & jest użyty do określenia miejsca, w którym ma nastąpić wyrównanie, \cr oznacza koniec linii. Nasze dane zostaną więc wydrukowane następująco:

$$\begin{aligned} x+x+2-2x &= 2x+2-2x \\ &= 2. \end{aligned}$$

Oczywiście podane tu możliwości są zaledwie maleńkim fragmentem tego, co można osiągnąć za pomocą TEXa. Nie pokazałem tu ani jak drukować całki czy równania różniczkowe, ani też w jaki sposób można drukować macierze. Ta notatka nie ma jednak na celu nauczenia posługiwania się TEXem. Ilustracje pokazują fragmenty prac złożonych za jego pomocą.

Definition 4. Transition combinators ([4,5])

$$S_{pn} = \lambda\alpha\beta a_1 \ldots a_p x_0 x_1 \ldots x_n . \alpha a_1 \ldots a_p (\beta a_1 \ldots a_p x_0 x_1 \ldots x_n)$$

Property 4. Distribution law

$$NA_p(NP_k(\alpha, \beta), \gamma) = S_{pk}(NA_p(\alpha, \gamma), NA_p(\beta, \gamma))$$

And by this property, we can define the target-machine instructions and function *distr* to distribute the symbol table information. These definitions are given in Table 4. Figure 3 shows the code for our example after distribution.

Property 10 .

For any $\iota \neq l_1$

$$\alpha \mapsto \mathrm{W}(\beta).\iota = \alpha \mapsto \mathrm{GOTO}[l_2] \bullet \mathrm{REF}[l_1] \bullet \beta \mapsto \mathrm{REF}[l_2] \bullet \mathrm{PASS}.\iota$$

Proof: **The only interesting case is when $\alpha \mapsto \theta_0.\iota\sigma = \theta_0.\iota'\sigma'$. Since α is without l_1, l_2, then $\iota' \neq l_1$ and $\iota' \neq l_2$. Let us consider two cases: $\iota' = nil$ and $\iota' \neq nil$.**

Case 1. **The lefthand side evaluates to (nil, σ'). The evaluation of the righthand side is:**

$$\begin{aligned}
&\mathrm{GOTO}[l_2] \bullet \mathrm{REF}[l_1] \bullet \beta \mapsto \mathrm{REF}[l_2] \bullet \mathrm{PASS}.nil\sigma' = \\
&\quad \mathrm{REF}[l_1] \bullet \beta \mapsto \mathrm{REF}[l_2] \bullet \mathrm{PASS}.l_2\sigma' = \\
&\quad \beta \mapsto \mathrm{REF}[l_2] \bullet \mathrm{PASS}.l_2\sigma' = \\
&\quad \mathrm{REF}[l_2] \bullet \mathrm{PASS}.l_2\sigma' = \\
&\quad \mathrm{PASS}.nil\sigma' = (nil, \sigma')
\end{aligned}$$

O TEXu napisano już sporo prac i podręczników. Do najważniejszych należy *The Joy of TEX* Michaela Spivaka [2] przeznaczona dla matematyków. Podzbiór systemu omawiany w tej pracy ma być standardem proponowanym dla prac zgłaszanych do American Mathematical Society. Jest to praca napisana przystępnie i umożliwia szybkie nauczenie się tych szczegółów, które są potrzebne do pisania prac matematycznych. Drugim bardzo ważnym podręcznikiem jest książka autora systemu, Donalda E. Knutha, *TEX and METAFONT. New Directions in Typeset-*

ting [1]. Ten podręcznik jest szczególnie polecany dla tych, którzy chcą dokładniej zrozumieć system oraz definiować własne, bardziej wyrafinowane symbole.

Na zakończenie kilka danych technicznych. Jak już wspomniałem, TEX jest dostępny na większości uniwersytetów amerykańskich. Ponieważ został zaprogramowany w Pascalu, więc jest dostępny na niemal wszystkich komputerach. Oczywiście, aby móc otrzymać wyniki na papierze, musimy podłączyć do komputera specjalne urządzenia wyjściowe. Najlepsze do tego celu są jeszcze ciągle dość drogie drukarki laserowe lub fotoskłady rastrowe. Drukarki laserowe działają na podobnej zasadzie jak kopiarki, to znaczy korzystają z matryc odtwarzalnych oraz specjalnego papieru i farby.

Gdy chcemy drukować pracę czy książkę w większym nakładzie, powinniśmy korzystać z maszyn do fotoskładu produkujących trwałe matryce (klisze fotograficzne). Do niektórych uproszczonych tekstów możliwe jest również wykorzystanie fotoskładu ze sztywno określonym alfabetem. W trakcie przygotowywania pracy bardzo użyteczne mogą być graficzne monitory ekranowe.

W Instytucie Informatyki Uniwersytetu Warszawskiego od około roku prowadzone są prace nad zainstalowaniem TEXa produkującego dane dla fotoskładu. Być może wkrótce pierwsze teksty składane przy jego użyciu będą dostępne również u nas. Digital Equipment Corporation nazywa system opracowany przez Knutha *najważniejszym wynalazkiem bieżącego stulecia w zakresie składania tekstów* (...) *ważnym prawie tak samo jak wprowadzenie prasy drukarskiej przez Gutenberga.*([3])

Podręczniki TEXa

[1] D. E. Knuth, *TEX and METAFONT. New Directions in Typesetting*, American Mathematical Society and Digital Press, Bedford, Massachusetts, 1979.

[2] M. Spivak, *The Joy of TEX. A Gourmet Guide to Typesetting Technical Text by Computer*, American Mathematical Society, Providence, Rhode Island, 1982.

Literatura uzupełniająca

[3] A. V. Aho, S. C. Johnson and J. D. Ullman, *Typesetting by ACM considered harmful*, Communications of the ACM 18, 1975.

([3]) Artykuł ten traktujemy jako przyczynek do dyskusji dotyczącej wprowadzenia w Polsce systemów takich jak TEX. Problem dotyczy zarówno sprzętu (analiza tego, czym już dysponujemy w kraju, a co trzeba kupić oraz kosztów bieżących), jak i szkolenia autorów czy całych środowisk. W przyszłości autor-matematyk, dysponujący możliwością wydruku lub wyświetlenia na ekranie próbek swej pracy, mógłby zastąpić zecera (nie wszyscy oczywiście autorzy zechcą, czy będą w stanie podjąć się tej dodatkowej pracy, ale entuzjastów będzie wielu). Należy też rozważyć, w jakim stopniu tą drogą będzie można przyspieszyć druk tekstów matematycznych (*przyp. Red.*).

[4] American Mathematical Society, *Development of the Photon for efficient mathematical composition*, Final report, May 10, 1965.

[5] American Mathematical Society, *Development of computer aids for tape-control of photocomposing machines*, Final report, Section B, August 1968; Section C, January 1969.

[6] M. P. Barnett, *Computer typesetting: Experiments and prospects*, M. I. T. Press, Cambridge, Massachusetts, 1965.

[7] P. J. M. Coueignoux, *Generation of roman printed fonts*, Ph. D. thesis, Dept. of Electrical Engineering, M.I.T., June, 1975.