

Zdzisław SROCZYŃSKI
Politechnika Śląska, Instytut Matematyki

PREZENTACJA DOKUMENTÓW INTERNETOWYCH ZAWIERAJĄCYCH ZŁOŻONĄ NOTACJĘ MATEMATYCZNĄ DLA POTRZEB OSÓB NIEWIDOMYCH

Streszczenie. W artykule przedstawiono problemy związane z zapewnieniem osobom niewidomym dostępu do dokumentów internetowych zawierających złożoną notację matematyczną. Przedstawiono implementację rozwiązania polegającego na przetwarzaniu specjalnie dostosowanych dokumentów na serwerze tak, by były one czytelne zarówno dla zwykłego odbiorcy, jak i możliwe do wydrukowania na drukarce brailowskiej. Druga część artykułu zawiera analizę możliwości przetwarzania już istniejących dokumentów matematycznych dla potrzeb osób niewidomych.

Słowa kluczowe: MathML, serwery WWW, przeglądarki internetowe, notacja matematyczna, język Braille'a, rozpoznawanie, niewidomi.

DEVELOPMENT OF WEB PAGES WITH COMPLEX MATHEMATICAL NOTATION FOR VISUALLY IMPAIRED PEOPLE

Summary. The paper describes main problems in presenting complex mathematical documents to blind people. An implementation of automatic translation from web math notation to brail printer format has been described. In the second part of the article possible solutions for machine translation of existing, printed math documents are considered.

Keywords: MathML, web servers, web browsers, mathematical notation, Braille language, recognition, visually impaired people, blind people.

1. Wprowadzenie

Publikacja dokumentów zawierających złożoną notację matematyczną w sieci Internet stanowi istotne wyzwanie dla autorów i wymaga szczególnego wysiłku od odbiorców. Opra-

cowane dotychczas standardy, takie jak oparty na XML język MathML [3], nie zdobyły szerokiego rzesz użytkowników głównie ze względu na trudności implementacyjne. W pracy [4] przedstawiono ideę pozwalającą na stworzenie serwera przetwarzającego publikacje matematyczne ze względu na rodzaj przeglądarki używanej do oglądania dokumentów. Rozwiązanie takie daje możliwość opracowania jednolitego i spójnego systemu publikacji, niezależnego od wyposażenia i umiejętności odbiorcy. Częścią takiego systemu powinny być również narzędzia edycyjne, takie jak np. opisane w pracy [7].

Jednym z wniosków narzucających się podczas testowania opisanego powyżej systemu była możliwość rozszerzenia translacji na języki opisu dokumentu stosowane przez osoby niewidome i niedowidzące, a więc matematyczny dialekt języka Braille'a. Okazało się, że specyfikacja tegoż języka podlega ciągłemu rozwojowi i jej polska kodyfikacja zakończyła się w zasadzie dopiero pod koniec 2002 roku. Dokument [8] zawiera przyjęte przez najważniejsze ośrodki krajowe zasady posługiwania się językiem Braille'a przy opracowywaniu dokumentów zawierających notację matematyczną, fizyczną i chemiczną.

W pierwszej części artykułu przedstawiono rozszerzenie translatora przetwarzającego dokumenty na serwerze [4] o zapis w formacie drukarki brailowskiej. Zmiana urządzenia wyjściowego pociąga za sobą przy tym wyłącznie konieczność dołączenia dodatkowej tabeli konwersji kodów używanych przez dane urządzenie, na przykład linijkę brailowską. Sam wybór drukarki brailowskiej podyktowany był dostępnością takiego urządzenia do testów.

Nie wszystkie jednak dokumenty da się przygotować do odpowiedniej obróbki na serwerze, ponieważ wymaga to znacznych nakładów pracy oraz przeszkolenia licznego personelu zaangażowanego w proces składu publikacji. Olbrzymia baza książek jest dostępna wyłącznie w wersji czarnodrukowej i nawet częściowa automatyzacja przepisywania tych pozycji byłaby bardzo cenna. Oczywiście problem ten nie dotyczy zwykłego tekstu, ponieważ współcześnie dostępne systemy Optical Character Recognition (OCR) pozwalają w wysokim stopniu zautomatyzować konwersję dokumentów do postaci elektronicznej. Jednak dla języka technicznego, a w szczególności zapisu matematycznego, zadanie to okazuje się znacznie bardziej złożone. Dotychczasowe próby rozwiązania tego problemu nie dały w pełni satysfakcjonujących rezultatów.

Stąd też druga część artykułu zawiera przegląd dotychczasowych prób zapewnienia automatycznego rozpoznania struktury wzorów matematycznych zapisanych w postaci graficznej. Zagadnienie to rzecz jasna nie ogranicza się wyłącznie do zapewnienia osobom niewidomym dostępu do dzieł technicznych i matematycznych. Można łatwo wskazać inne zastosowania, jak udostępnianie drukowanych prac w nauczaniu na odległość, analiza ręcznie pisanych wyrażeń matematycznych w aplikacjach edukacyjnych i urządzeniach przenośnych, interaktywne systemy dydaktyczne typu tablica elektroniczna czy też przetwarzanie dotychczas dostępnych

stron WWW zawierających notację matematyczną zapisaną w postaci plików graficznych w celu katalogowania i wyszukiwania informacji. Jednak zapewnienie osobom niewidomym i niedowidzącym sposobu automatycznego przetwarzania dokumentów czarnodrukowych na format akceptowany przez urządzenia drukujące i przedstawiające znaki alfabetu Braille'a jest z pewnością bardzo istotnym wyzwaniem.

2. Translacja notacji matematycznej w języku MathML na format używany w matematycznym dialekcie języka Braille'a

Największym problemem przy opracowaniu translatora pozwalającego przetwarzać wyrażenia matematyczne zapisane w języku MathML na punktowy zapis w języku Braille'a okazała się specyfika notacji matematycznej używanej przez osoby niewidome. Wynika ona po części z dążenia do jak największego skrócenia zapisu, a po części z genezy zapisu matematycznego Braille'a, który jest standardem *de facto* i do niedawna w Polsce nie istniała nawet jednolita jego specyfikacja. Pismo Braille'a ma charakter liniowy, więc przekształcanie powszechnie spotykanych w matematyce zapisów wykorzystujących przestrzenny rozkład pewnych symboli napotyka na pewne trudności. Niekiedy jednak zapis brajlowski powstaje poprzez analogię do wyglądu zapisu czarnodrukowego. Ma to miejsce na przykład dla ułamków zwykłych, których mianownik zapisuje się zwykle na tzw. obniżonym poziomie, co odaje wizualny aspekt zapisu matematycznego.

Pismo Braille'a jest ograniczone do 64 znaków, stąd konieczność wykorzystywania tych samych znaków w kombinacji ze znakami specjalnymi do kodowania różnych symboli. Znaczenie matematycznego zapisu Braille'a zależy więc silnie od kontekstu.

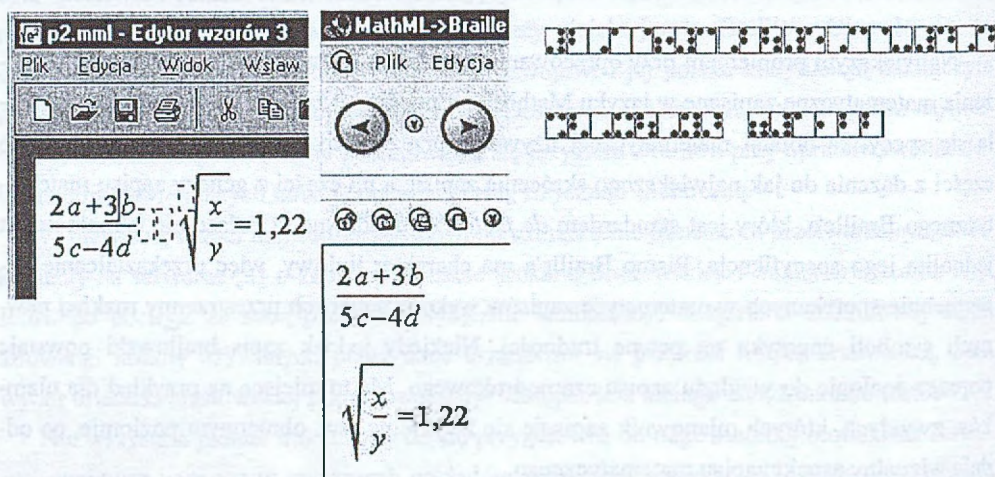
Graficzna notacja matematyczna pozwala wykorzystać zmieniające się rozmiary pewnych symboli, np. kreski ułamkowej, znaku pierwiastka lub zmieniającą się pozycję symboli, np. zmianę poziomu wskaźników do zapewnienia właściwego grupowania wyrażeń matematycznych. W zapisie Braille'a uzyskanie podobnych efektów wymaga przekształcania wyrażeń z użyciem nawiasów lub stosowania specjalnych konwencji, takich jak tzw. znak ciągłości czy znaki początku i końca ułamka.

Podsumowując należy stwierdzić, że zapis Braille'a nie umożliwia zróżnicowania informacji poprzez zmianę wielkości, kształtu i wzajemnego położenia znaków [8].

W ramach testów opracowano translator jako moduł eksportu danych w programie Edytor wzorów [7] oraz jako moduł rozszerzający serwer WWW w sposób opisany w pracy [4]. Zbudowanie takiego systemu daje realną szansę publikowania dokumentów dostępnych powszechnie w Internecie, a przy tym czytelnych dla osób niewidomych bez konieczności ich specjalnego uprzedniego przetwarzania.

Obecnie translator pozwala tworzyć i konwertować proste wyrażenia algebraiczne zawierające:

- duże i małe litery,
- liczby całkowite i zmiennoprzecinkowe,
- podstawowe operatory,
- nawiasy z uwzględnieniem rodzaju,
- ułamki zapisywane z użyciem znaku ciągłości,
- pierwiastki kwadratowe.



Rys. 1. Dokument zawierający notację matematyczną w trakcie edycji w programie Edytor wzorów (z lewej). Ten sam dokument podczas wizualizacji w przeglądarce Opera po przetworzeniu przez translator na serwerze WWW (po środku). Wydruk dokumentu z drukarki brajlowskiej (z prawej; symulacja za pomocą czcionki brajlowskiej)

Fig. 1. Document with embedded math notation during edition (on the left), visualized on-the-fly by the web server extension (in the middle in Opera browser) and printed with Braille-printer (simulated using Braille font)

Wprowadzenie w życie opisanego rozwiązania wymaga rozróżnienia przeglądarki internetowej wspomagającej pracę osób niewidomych. Identyfikator przeglądarki przekazywany jest do serwera w zmiennej HTTP_USER_AGENT. Dopisanie specyficznego zapisu do ciągu identyfikującego program Internet Explorer firmy Microsoft wymaga jedynie niewielkiego uzupełnienia w rejestrze systemowym Windows. Niektóre inne przeglądarki, jak na przykład Konqueror dostępny w KDE systemu Linux pozwalają na ustawianie tego identyfikatora bezpośrednio w okienku ustawień programu. Można przyjąć, że przeglądarka z identyfikatorem zawierającym na dowolnej pozycji ciąg znaków "Braille" powinna być przez serwer rozpo-

znawana jako wymagająca przetwarzania zapisu w języku MathML na odpowiedni kod w notacji Braille'a. Oczywiście w pełnej translacji należy brać pod uwagę nie tylko przekształcanie wyrażen matematycznych, ale również zwykłego tekstu literackiego. Zadanie to jednak łatwo rozwiązać za pomocą tabeli konwersji.

Można również stworzyć specjalną przeglądarkę dla osób niewidomych. Jej niepodważalną zaletą mogłaby być pełna automatyzacja procesu przesyłania zakodowanej w języku Braille'a strony WWW do urządzenia wyjściowego. Z drugiej jednak strony wiele dokumentów internetowych, nie zawierających złożonych zapisów matematycznych, łatwiej przedstawiać przy wykorzystaniu standardowych przeglądarek rozszerzonych o np. możliwość syntezy mowy.

Otrzymanie ostatecznego efektu działania opisanego powyżej systemu wymaga rzecz jasna wydrukowania otrzymanej przetworzonej strony WWW na drukarce brajlowskiej. Łatwość wykonania tego zadania zależy w dużym stopniu od sposobu komunikacji drukarki z używanym systemem operacyjnym.

3. Rozpoznanie struktury wzorów matematycznych zapisanych w postaci graficznej

Notacja matematyczna przez długi czas nie doczekała się powszechnie obowiązującego standardu zapisu. System TeX/LaTeX zdobył stosunkowo dużą popularność w środowiskach akademickich, jednak wielu użytkowników używa innych narzędzi, opierających się na wizualnych edytorach. Złożone dokumenty zawierające notację matematyczną są więc dostępne w wielu odmiennych formatach, a znaczna ich część to dokumenty w postaci drukowanej oraz strony internetowe, na których wzory zapisano w postaci plików graficznych.

Taki stan rzeczy w znaczący sposób utrudnia katalogowanie, udostępnianie i wyszukiwanie dokumentów matematycznych. Nie sposób bowiem zapewnić jednakową obsługę dokumentów o różnym charakterze. Szczególnie kłopotliwe są druki, ponieważ powszechnie dostępne oprogramowanie do rozpoznawania pisma nie obsługuje wzorów matematycznych.

Jest to szczególnie dotkliwe dla osób niewidomych, które praktycznie pozbawione są możliwości odbioru dokumentów matematycznych za pomocą metod automatycznych. Opracowanie metod zapewniających właściwą prezentację tworzonych właśnie dokumentów, jak to opisano w pkt. 2, rozwiązuje problem połowicznie, gdyż ciągle poza zasięgiem niewidomych pozostaje olbrzymia baza dawniej powstałych dokumentów.

Jedynie opracowanie metody zapewniającej możliwość automatycznego rozpoznawania struktury wzorów matematycznych zapisanych w postaci rastrowej mapy bitowej może udo-

stąpić kompleksowo zbiór internetowych dokumentów matematycznych osobom niewidomym oraz jako materiał do dalszego przetwarzania przez różnego rodzaju roboty internetowe, jak np. wyszukiwarki.

Notacja matematyczna jest rodzajem języka wizualnego, który można zdefiniować poprzez zbiór diagramów uznawanych za poprawne wyrażenia w tym języku. Właściwości syntaktyczne i semantyczne takiego języka określają przestrzenne zależności pomiędzy symbolami w diagramie [2]. Jednym z istotnych zadań jest więc zdefiniowanie modelu danych pozwalającego przechowywać i przetwarzać wyrażenia matematyczne.

Równie istotny problem do rozwiązania to opracowanie metod przetwarzania symboli i ich atrybutów w celu wyodrębnienia kolejnych fragmentów przetwarzanego wyrażenia. Analiza struktury wyrażenia matematycznego nie jest zadaniem trywialnym ze względu na dużą ilość obowiązujących reguł zapisu, często wynikających z tradycji. Zaproponowane dotychczas podejścia często charakteryzują się ograniczeniami wynikającymi z poczynionych założeń [10][11] lub dużą złożonością obliczeniową [1]. Można je podzielić na podstawowe kategorie:

- metody syntaktyczne,
- metody rekursywnego podziału na podwyrażenia (projection-profile cutting),
- metody translacji grafu,
- metody proceduralne, bazujące na analizie geometrycznych zależności między obiektami w badanym wyrażeniu, szczególnie przydatne w systemach pracujących on-line (przyrostowo) [10].

Niestety, niektóre z podjętych dotychczas projektów w praktyce nie wyszły poza rozpoznawanie prostych wyrażeń wielomianowych, ułamków i funkcji trygonometrycznych. Kłopoty sprawiały szczególnie wzory zagnieżdżone, takie jak pierwiastki i macierze [10].

Etapy rozpoznania wyrażenia matematycznego można podzielić na [11]:

- akwizycję mapy bitowej,
- wyodrębnienie listy symboli wraz z ich atrybutami,
- analizę struktury,
- analizę znaczenia wzoru.

Większość autorów dotychczasowych opracowań [1][11] zgodna jest co do tego, iż rozwiązanie zadania rozpoznania wzoru matematycznego zapisanego w postaci graficznej powinno skupiać się na opracowaniu metod analizy strukturalnej. Rozpoznawanie pojedynczych symboli jest bowiem dziedziną dobrze rozwiniętą i udokumentowaną, a praktyczne rezultaty (dostępne również w komercyjnych aplikacjach OCR) są co najmniej zadowalające. Jednocześnie poznanie znaczenia wzoru nie jest konieczne do przetłumaczenia

go na język Braille'a, choć w niektórych przypadkach (potęgi, indeksy) może być przydatne ze względu na kontekstowość zapisu.

Artykuł [1] zawiera propozycję analizy dwuwymiarowej notacji matematycznej poprzez zastosowanie reguł określonej gramatyki kontekstowej podczas translacji grafu opisującego zależności między elementami notacji. Wierzchołki takiego grafu zawierają informacje o symbolu i jego charakterze, a krawędzie dane o względnym położeniu wierzchołków. Główne problemy napotkane w tym podejściu polegają na konieczności usuwania wieloznaczności wynikających z zastosowania różnych reguł gramatyki do tych samych części analizowanego grafu. Samo generowanie początkowego grafu, zawierającego informacje o charakterze i wzajemnych relacjach geometrycznych rozpoznanych symboli, również wymaga zapewnienia pewnych jego cech trudnych do zdefiniowania. Graf zawierający zbyt dużą ilość informacji o wzajemnych relacjach może bowiem generować więcej niż jedno wyrażenie wynikowe. Graf o zbyt małej liczbie krawędzi może zaś nie zawierać wystarczającej ilości informacji potrzebnych do odtworzenia pełnej struktury wzoru. Konieczność unikania tworzenia fałszywych krawędzi wymaga zastosowania dodatkowego algorytmu analizującego wzajemne położenie symboli, wprowadzającego podział sąsiedztwa geometrycznego na pewną dyskretną liczbę kategorii. W rozwiązaniu każdego z opisanych powyżej problemów istotną rolę pełnią reguły o charakterze heurystycznym.

Geometryczna struktura wzoru matematycznego może być również przedstawiona jako drzewo zawierające informacje o wzajemnych relacjach linii bazowej poszczególnych części wzoru [11]. Drzewo takie posiada następujące własności:

- grupowanie symboli oraz kierunek zapisu są jednoznaczne,
- wszelkie zależności geometryczne są jednoznaczne,
- przy użyciu nawiasów można otrzymać zlinearyzowany ciąg odpowiadający rozpoznanemu wzorowi wędrując przez liście drzewa,
- może zostać przekształcone w celu opisanego bardziej złożonych wyrażeń oraz wyeliminowania błędów wynikających z niejednoznaczności danych o symbolach.

Bazujący na tym podejściu system wymaga doprecyzowania funkcji badających geometryczne zależności między symbolami w indeksach, zapewnienia grupowania rozpoznanych symboli (np. cyfr w liczby, liter w nazwy funkcji). Macierze i wzory wielowierszowe nie są rozpoznawane, pewne problemy sprawiają też symbole o niestandardowych rozmiarach.

Dopiero połączenie rozwiązania opisanego w punkcie drugim z praktyczną i efektywną implementacją algorytmów, których przegląd zawarto powyżej, dałoby możliwość stworzenia kompleksowego systemu udostępniającego bibliotekę publikacji matematycznych na potrzeby osób niewidomych.

Opracowane dotąd aplikacje i moduły rozszerzające funkcjonalność serwera WWW zapewniają proste i funkcjonalne metody opracowywania i udostępniania jedynie tych dokumentów, które tworzone są lub przepisane na bieżąco. Jest to jednak solidna podstawa dla dalszego rozwoju opisywanych rozwiązań.

4. Podsumowanie

W artykule przedstawiono implementację rozszerzenia funkcjonalności serwera WWW o automatyczną translację notacji MathML na zapis w języku Braille'a. Rozwiązanie to pozwala udostępnić osobom niewidomym bibliotekę dokumentów matematycznych i technicznych pod warunkiem, że zostaną one zapisane lub przepisane przy użyciu standardu MathML. Dokumenty takie mogą być przeglądane bezpośrednio w pewnych przeglądarkach internetowych, przetwarzane wstępnie do formatu graficznego na serwerze dla innych (nie obsługujących MathML bezpośrednio) lub konwertowane do notacji Braille'a dla przeglądarek specjalnie oznaczonych.

Udostępnienie w podobny sposób bazy matematycznych dokumentów czarnodrukowych, dla których nie istnieją wersje elektroniczne, wymaga opracowania metod rozpoznawania struktury wzorów matematycznych zapisanych w formacie graficznym. Druga część artykułu zawiera przegląd dotychczas proponowanych rozwiązań w tej dziedzinie. Niestety, niedoskonałości i ograniczenia założeń nie pozwalają jak dotąd na efektywne wykorzystanie ich implementacji w praktyce.

LITERATURA

1. Lavirotte S., Pottier L.: *Mathematical Formula Recognition Using Graph Grammar*. Electronic Imaging 1998.
2. Marriot K., Meyer B., Wittenburg K. D.: *A Survey of Visual Language Specification and Recognition*. Visual Language Theory, str. 5-85, Springer-Verlag, New York 1998.
3. *Mathematical Markup Language (MathML) 2.0 Working Draft*. <http://www.w3.org/>.
4. Sroczyński Z.: Serwer WWW jako serwer dokumentów zawierających złożoną notację matematyczną. *Studia Informatica* Vol. 23 No. 2B (49), Gliwice 2002.
5. Sroczyński Z.: MathML — język opisu wyrażeń matematycznych w dokumentach internetowych. *Studia Informatica* Vol. 21 No. 3 (41), Gliwice 2000.

6. Sroczyński Z.: Rozszerzanie możliwości przeglądarek internetowych na przykładzie modułów plug-in dla Netscape Navigator. *Studia Informatica* Vol. 21 No. 3 (41), Gliwice 2000.
7. Sroczyński Z.: Edytor wzorów dla systemu LaTeX. Praca dyplomowa magisterska. Instytut Informatyki Politechniki Śląskiej, Gliwice 1997.
8. Świerczek J. (red.): Brajlowska notacja matematyczna, fizyczna, chemiczna. Kraków 2002.
9. Świerczek J.: Brajlowska notacja matematyczna. Opracowanie Specjalnego Ośrodka Szkolno-Wychowawczego dla Dzieci Niewidomych i Słabo Widzących. Kraków 1998.
10. Wan B., Watt S., An Interactive Mathematical Handwriting Recognizer for the Pocket PC. *MathML Conference* 2002.
11. Zanibbi R.: Recognition of Mathematics Notation via Computer Using Baseline Structure. External Technical Report. Queen's University, Kingston, Ontario, Canada 2000.

Recenzent: Prof. dr hab. inż. Tadeusz Czachórski

Wpłynęło do Redakcji 11 kwietnia 2003 r.

Abstract

The paper describes main problems in presenting complex math documents to blind people. An implementation of automatic translation from web math notation (MathML) to Braille printer format has been described. The results are shown at Fig. 1. Basic problems observed due to the nature of Braille notation were discussed. The system at the moment does not cover more complicated expressions, but it is still under development.

In the second part of the article possible solutions for machine translation of existing, printed math documents are considered. The main methods of automatic recognition of the mathematical expressions saved in bit-mapped formats are considered. As a conclusion a lack of practically useful and efficient implementations is pointed.

Adres

Zdzisław SROCZYŃSKI, Politechnika Śląska, Instytut Matematyki, ul. Kaszubska 23,
44-101 Gliwice, Polska