

МЕТОД ОЗВУЧЕННЯ МАТЕМАТИЧНИХ ФОРМУЛ ТА СИМВОЛІВ УКРАЇНСЬКОЮ МОВОЮ

У статті досліджено проблему автоматичного озвучення математичних формул та символів. Авторами отримані практичні результати, які дають змогу розробити достатньо якісну програму озвучення формул українською мовою. Авторами розроблено інформаційну технологію, яка складається з набору методів обробки і перетворення для автоматичного озвучення математичних формул та символів різних форматів українською мовою.

Ключові слова: математичні формули, символи, синтезатор мови, метод озвучення, перетворення, книга, що «розмовляє», MathML, TeX, DAISY.

В данной статье исследуется проблема автоматического озвучивания математических формул и символов. Авторами получены практические результаты, позволяющие разработать достаточно качественную программу озвучивания формул на украинском языке. Авторами разработана информационная технология, которая состоит из набора методов обработки и преобразования для автоматического озвучивания математических формул и символов различных форматов на украинском языке.

Ключевые слова: математические формулы, символы, синтезатор речи, метод озвучивания, преобразование, книга, которая «разговаривает», MathML, TeX, DAISY.

This paper describes problem of automatic speech mathematical formulas and symbols in Ukrainian. The authors obtained practical results that allow create the quality automatic mathematical speech system. The authors developed information technology that consists of a set of processing and conversion methods for automatic reading mathematical formulas and symbols of the Ukrainian language, which written in a variety of formats.

Key words: mathematical formulas, equations, speech synthesizer, method to automatic reading, conversion, digital talking book, TTS, MathML, TeX, DAISY.

Вступ. У цілому світі є чимало педагогічних, програмних і технічних засобів для полегшення процесу навчання та комунікації людей із вадами зору, проте подальший розвиток у цьому напрямі неможливий без комп'ютерного подання та адаптації навчально-методичного матеріалу до потреб людей із вадами зору. Нині проблемою навчання осіб із вадами зору (дошкільнят, школярів, студентів, осіб, що бажають підвищити кваліфікацію) займаються багато вчених, дослідників, лікарів у багатьох країнах світу. В Україні ця проблема залишається досить проблематичною і болючою. Її вирішення дозволить суттєво спростити процес доступу до інформації незрячим користувачем. Чималий вклад у створення новітніх інформаційних технологій цього профілю зроблено у Швеції, Японії, Німеччині, США, Франції, Канаді, Данії, Англії та інших провідних країнах.

Формулювання проблем незрячих охоплює велику кількість досліджень, зокрема на теми програмного й апаратного забезпечення, розробки методів та засобів реалізації мовного синтезу й розпізнавання мови.

Аналіз досліджень і публікацій. На сьогодні процес озвучення формул та символів українською

мовою залишається не дослідженим, у результаті чого не має можливості цілісно озвучити контент для україномовних користувачів із вадами зору. Хоча і відомі засоби озвучення математичних формул англійською мовою [1; 2], проте вони не вирішують проблему для інших мов.

Останні дослідження в галузі озвучення математичних формул [3; 4; 5] зосереджені лише на озвученні формул англійською мовою. Дослідження цих засобів показало, що вони не придатні для озвучення формул українською мовою та не містять засобів локалізації.

Задача озвучення математичної формули лежить у площині задач комп'ютерного перекладу, оскільки має усі ознаки таких задач, а саме: необхідність часткового розуміння змісту формули для правильної інтерпретації, побудову проміжного комп'ютерного запису формули та розроблення засобів кодування формули вихідною мовою.

Застосування популярного статистичного методу комп'ютерного перекладу [6] до читання формул не є доцільним, оскільки немає відомих баз паралельних текстів «формула-україномовний текст».

Задача читання формули найбільш подібна до задачі перекладу з використанням універсальної проміжної мови «interlingua» [7]. У випадку математичної формули такою проміжною мовою може бути мова математичної розмітки MathML, мова TeX, OpenOffice Math або інша.

Оскільки це сприяло прийняттю єдиного стандарту, до якого можуть конвертувати формули різні математичні редактори, таким стандартом стала мова математичної розмітки MathML (Mathematical Markup Language), рекомендована консорціумом W3C. Мова MathML базується на мові розмітки XML (eXtensible

Markup Language – розширювана мова розмітки) та призначена, насамперед, для комп'ютерного подання формул. Формула, представлена засобами мови MathML, подана деревовидною структурою, яка дає змогу однозначно описати семантику математичного виразу в інформаційній системі [8]. Кожен вузол дерева відповідає певній схемі компонування, а його гілки або нащадки відповідають підвиразам. Іншими словами, це графічне представлення математичної формули, яке показує, як теги MathML повинні бути вкладені один в один для правильного відображення математичного виразу на екрані (рис. 1).

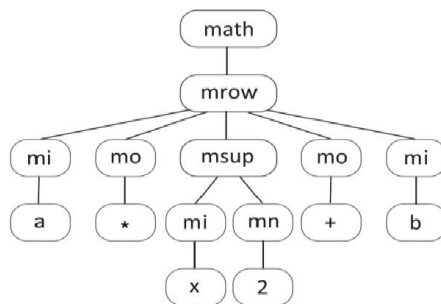


Рис. 1. Дерево MathML для формули $a * x^2 + b$

MathML використовує два способи кодування та запису математичних формул: презентаційний та семантичний. Презентаційний запис описує математичну символіку з формулами, які будуються з використанням схем виведення таких, як дроби, верхні та нижні індекси, тобто він описує візуальну форму представлення математичної формули, дає змогу керувати її виглядом, його розміром, кольором, символами, точним положенням кожного елемента. Семантичний запис відповідає за семантику формули, тобто, описує математичні об'єкти і функції, де для кожного вузла конструюється дерево обчислення згідно з деякою конкретною схемою, а гілки цього дерева відповідають підвиразам [9]. Обидва варіанти розмітки підпорядковуються одному базовому набору правил, проте, відрізняються тегами при записі формули. В одному документі MathML можливе поєднання одночасно двох варіантів запису, що необхідно врахувати при озвученні формули [10].

Альтернативним способом подання семантики математичної формули є формат OpenMath [11]. Формат OpenMath задає лише семантику математичної формули, натомість формат MathML може визначити структуру відображення математичної формули. Структура опису математичних об'єктів у OpenMath дає змогу використовувати цей формат у системах формальної логіки й у системах пошуку доведень. Як подальший розвиток MathML і OpenMath зараз створюється стандарт «OpenMathematicalDocuments» (OMDoc).

На сьогодні формат OpenMath ще не набув такої популярності, як MathML, а тому бракує засобів перетворення формул у цей формат.

Ще одним форматом, який класично використовується для редагування та видання наукової літератури є формат TeX. TeX – це система підготовки документів типографської якості, особливо сильною стороною

якої є набір математичних формул. Ця система була розроблена у 1970-х роках американським математиком Дональдом Кнудом.

TeX розглядається як спеціальна мова програмування: вихідний файл – це програма, що описує, як повинні бути розташовані на друкованій сторінці текст і формули, сама система TeX відіграє роль компілятора, а в результаті компіляції створюється документ типографської якості у форматі PDF, PostScript або в традиційному для TeX'a форматі DVI (DeVice Independent).

TeX, на відміну від мови MathML, не підтримує деревовидне подання формули та допускає неоднозначне трактування семантики математичного виразу.

Серед засобів перетворення документу TeX на мову MathML найбільш якісною вважається програма TtM, розроблена А. Хатчінсоном (I. Hutchinson) [12]. Ця програма перетворює документ TeX у файл формату XHTML, до складу якого входять MathML формули.

Аналіз документів, у яких зустрічаються математичні формули у форматі PDF, як і раніше є складним, невіршеним завданням. На сьогодні розпізнавання складних математичних формул із зображень можливе лише зі значними неточностями передавання змісту формули. Є два основних підходи, які частково вирішують цю проблему.

Перший підхід, який реалізований у програмі Infty [13], зводить проблему до традиційного підходу роботи з відсканованими документами. Він використовує технологію оптичного розпізнавання символів OCR (Optical character recognition) і застосовує структурний аналіз до отриманого результату. У той час як ця технологія розпізнає складні математичні формули в документі, вона ігнорує іншу текстову інформацію, що міститься в документах PDF.

Другий підхід використовує як графічну, так і текстову інформацію, яка міститься в документі PDF. Для цього використовується розроблений авторами дворівневий аналізатор контенту і структури математичних формул [14].

Використання програми Infty доцільне при розпізнаванні математичних формул із растрових зображень і PDF файлів та для подальшого запису мовою MathML.

Інші комерційні засоби редагування формул такі як MathType, MathCad, Mathematica та інші підтримують збереження формул мовою MathML. Для запису формул мовою MathML, які містяться в документах Microsoft Word та OpenOffice, використовують засоби перетворення документів у TeX, наприклад, GringEQ Word-to-LaTeX та writer2latex.

Найбільш потужним програмним засобом, який може конвертувати і озвучувати математичні формули є комерційний продукт MathDaisy [15]. MathDaisy озвучує формули лише англійською мовою, що є недоліком для українськомовних користувачів. MathDaisy забезпечує повну підтримку математики у книгах, що «розмовляють» у форматі DAISY.

Система керування цифровим доступом DAISY (Digital Accessible Information System) – це відкритий міжнародний стандарт для мультимедійного доступу, розроблений DAISY-консорціумом разом із працівниками бібліотек для осіб із фізичними недоліками та вадами зору у співпраці з передовими науковцями і користувачами з усього світу [16]. Технологія DAISY базується на синхронізації тексту, графіки і аудіо-записів на базі рекомендацій W3C (World Wide Web Consortium) відповідно до потреб людей, які вимагають особливого доступу до інформації.

Перевагою книги, що «розмовляє», поданої у форматі DAISY над звичайною аудіо-книгою є можливість індексування і пошуку інформації за ключовими словами. Користувачі можуть прослухати всю книгу лінійно, від початку до кінця, або використовуючи засоби навігації, які дозволяють швидко переходити від одного розділу до іншого, від абзацу до абзацу, або відзначати закладкою відповідні місця в тексті для повторного прослуховування. Існує два способи, які лежать в основі створення книги, що «розмовляє», а саме начитування контенту книги диктором у спеціально обладнаній студії звукозапису і озвучення контенту синтезатором мови.

Значний внесок у розроблення математичних моделей та методів для озвучення контенту зробив Алан В. Блек, доцент інституту мовних технологій Університету Карнегі-Меллона (Associate Professor, Language Technologies Institute, Carnegie Mellon University). Цілями його досліджень є практична реалізація обчислювальної теорії голосу і мови та створення комп'ютерного синтезу мови, максимально наближеної до природної, такої ж гнучкої та ефективної, як людська мова. Результатами досліджень стали сотні публікацій і монографій за тематикою синтезу і розпізнавання мови та розробки програмних комплексів, зокрема і Festival Speech Synthesis System.

В Україні проблемами синтезу мови займався відділ розпізнавання та синтезу звукових образів Міжнародного науково-навчального центру інформа-

ційних технологій та систем (МННЦІТС), м. Київ та Українська асоціація з оброблення інформації та розпізнавання образів (УАСОІРО) при ньому, Донецький Державний інститут штучного інтелекту (ДІШІ), в якому ведуться роботи з розпізнавання мовлення, а також велася робота в невеликих наукових кафедральних групах різних університетів.

Цілі статті. Дослідження, які було проведено, показали, що найчастіше проблема розуміння математичних формул виникає при читанні Інтернет сторінок із технічно-прикладним вмістом, при читанні фахових статей у форматі PDF та при роботі з друкованою книгою із використанням читаючого сканера. Це зумовлено відсутністю засобів, які б дозволяли під час роботи в мережі Інтернет у реальному часі перекладати та озвучувати математичні формули та символи, недосконалістю засобів розпізнавання математичних формул, представлених у вигляді рисунків і транслювання їх мовою MathML, відсутністю засобів озвучення математичного контенту, друкованого на папері.

Отже, задача озвучення формули зводиться до розроблення системи правил, за якими перетворюється внутрішнє представлення формули до зовнішнього представлення українською мовою.

Авторам статті не відомі публікації, де були б зібрані правила читання математичних формул українською мовою, проте відомі такі публікації російською мовою [17]. Ці публікації взяті авторами за основу правил читання, враховуючи схожість української та російської мов.

Процес аудіо-відтворення формул та математичних символів українською мовою не досліджений. Немає систем озвучення формул українською мовою, що значною мірою звужує можливості незрячої людини при навчанні прикладним предметам у навчальних закладах.

У статті обґрунтовано вибір мови MathML для внутрішнього подання формул та описано розроблений авторами метод озвучення математичних формул українською мовою.

Розв'язано задачу опрацювання, перетворення та озвучення математичних формул, які зустрічаються у текстах, з якими працює незрячий користувач. Розроблений метод є частиною інформаційної технології озвучення формул, записаних у різних форматах. Інформаційна технологія складається з набору правил, що дає змогу в автоматизованому режимі розібрати математичну формулу, подану в різних форматах (TeX, PDF, JPG), та озвучити її українською мовою за допомогою синтезатор мови, який називається «український голос UkrVox – Igor».

Розроблення інформаційної технології виконувались із врахуванням подальшої практичної реалізації автоматизованої прикладної програмної системи комп'ютерного опрацювання та подання даних для людей із вадами зору. Результати досліджень авторів статті містять розроблену програму перетворення та озвучення математичних формул українською мовою, що вирішує проблему повноцінного вивчення математики для незрячої людини та робить доступними електронні книги, в яких використовуються математичні символи.

Вдосконалення розроблених технологій може здійснити значний внесок у вирішення проблеми інтеграції людей з вадами зору у суспільство, зокрема, навчання сліпих та людей із частковою втратою зору у вищих навчальних закладах.

Виклад основного матеріалу. Для озвучення математичних формул, записаних у різних форматах, авторами розроблено інформаційну технологію, яка складається із сукупності методів опрацювання та перетворення формул. Розроблена інформаційна технологія містить чотири етапи опрацювання математичної формули:

1) **Сканування формули.** На цьому етапі відбувається сканування формули з відповідного інформаційного джерела (газета, книга, електронний документ, зображення тощо).

2) **Розпізнавання та перетворення формули до MathML.** Наступним етапом є розпізнавання відсканованої формули, а також, подання її мовою MathML. Розпізнавання формули або графічного зображення відбувається засобами програми Infty, яка дає змогу перекласти формулу мовою MathML або TeX. У результаті ми отримуємо файл із розширенням *.mml, в якому записана формула.

3) **Формування тексту для озвучення.** На цьому етапі за допомогою розроблених авторами правил відбувається розбір математичної формули на ідентифікатори, числа та оператори для подальшого озвучення. На виході отримуємо текстовий файл у форматі TXT: «ікс квадрат мінус чотири дорівнює ка квадрат».

4) **Озвучення або збереження у форматі WAV.** Тут виконується озвучення текстового файлу, який був створений на попередньому етапі. Підключається синтезатор мови, в якому використовується український голос. У результаті озвучення отримуємо звуковий файл у форматі WAV.

Для озвучення сформованого текстового файлу математичної формули (етап 4) використано синтезатор мови, який називаються «український голос UkrVox – Ігор», за основу якого взято голос диктора українського радіо Ігоря Мурашко. Цим голосом можна озвучувати українські (а також англійські та російські) тексти. «Український голос UkrVox – Ігор» має сильну лінгвістичну базу, великий словник (кілька десятків тисяч слів), підтримку словотворення та морфологічного аналізу текстів [19].

Результатом озвучення є аудіо-файл із записом формули у форматі *.wav.

Аналіз результатів та висновки.

На рис. 2 схематично зображено сукупність методів перетворення формул із різних форматів до мови MathML для подальшого озвучення. Ці методи формують інформаційну технологію перетворення формул для їх озвучення або запису в аудіо-файл.

Спочатку відбувається аналіз формату, в якому подана математична формула і приймається рішення, яким чином вона буде перетворена до мови MathML. Після запису мовою MathML, формула перетворюється в текст і озвучується, відповідно до алгоритму, описаного вище.

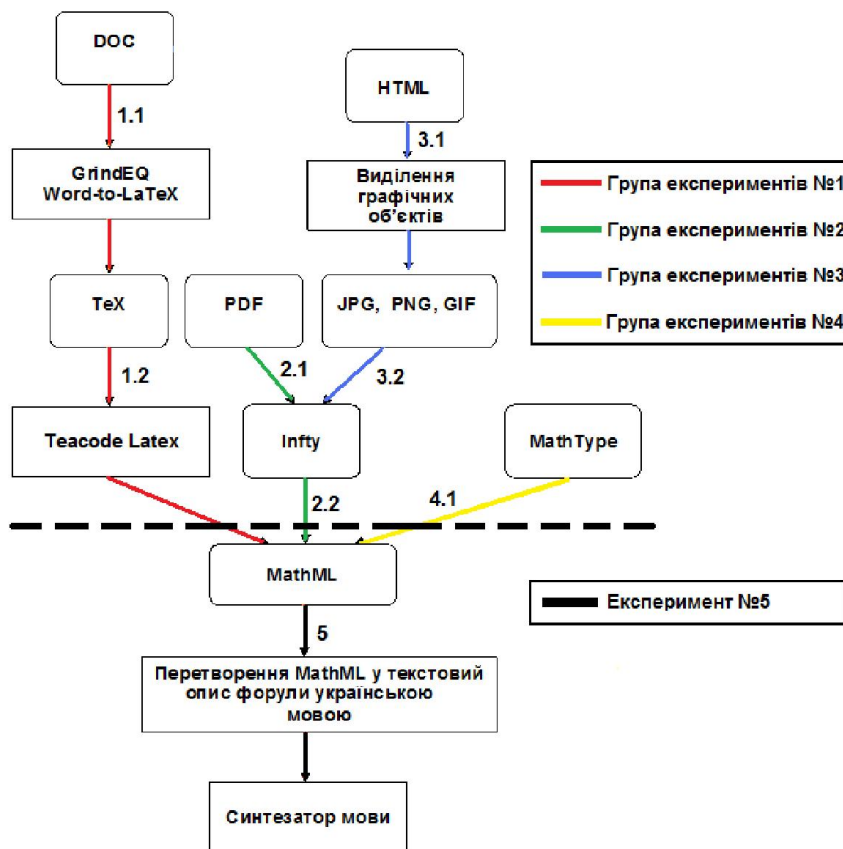


Рис. 2. Інформаційна технологія озвучення формул

Проведено чотири групи експериментів щодо перетворення математичних формул у різних форматах на мову математичної розмітки MathML, а також, експеримент з озвучення отриманих MathML формул (Експеримент № 5).

Група експериментів № 1 (*.doc → «GrindEQ Math» → TeX → Teacode Latex → MathML). Перша група експериментів полягає в перетворенні за допомогою плагіну GrindEQ Math Utilities та он-лайн сервісу Teacode Latex на мову MathML набору формул, записаних у форматі текстового документу Microsoft Word.

Група експериментів № 2 (*.pdf → Infty → MathML). Друга група експериментів полягає в перетворенні на мову MathML набору формул, записаних у форматі Adobe Reader, *.pdf за допомогою програми для розпізнавання символів Infty.

Група експериментів № 3 (*.html → виділення графічних об'єктів → JPG, PNG, GIF → Infty → MathML). Третя група експериментів базується на перетворенні набору формул, записаних у вигляді веб-сторінки *.html на мову MathML засобами програми Infty.

Група експериментів № 4 (MathType → MathML). Четверта група експериментів полягає в перетворенні

на мову MathML набору формул, записаних засобами MathType.

Експеримент № 5 (озвучення українським голосом формули, яка записана мовою MathML).

Для проведення експерименту взято 150 формул. Всі формули поділено на групи: рівняння (раціональні, квадратні, лінійні), нерівності, тотожні вирази, дробові вирази та функції. У кожній групі присутні формули різної складності. Складність формули визначалась кількістю операндів, чисел та математичних операцій (проста формула, формула середньої складності, складна формула).

Експеримент полягає в коректному сприйнятті на слух озвученої математичної формули та записі її у блокноті після/під час прослуховування. Озвучення групи математичних формул відбувається за допомогою розробленої авторами програми MathPlay.

Для проведення експерименту було запрошено групу з п'яти зрячих студентів спеціальності комп'ютерні технології. Студентам було видано блокноти та ручки для запису прослуханих формул.

Кожна із 150 спроб проведеного експерименту вважалась успішною, якщо після двох прослуховувань студент записав формулу в блокноті без помилок.

Результати експерименту зведені в табл. 1.

Таблиця 1

Результати експерименту № 5

Вірних спроб Учасники експерименту	прості формули	формули середньої складності	складні формули
1-й студент	50	50	48
2-й студент	49	50	44
3-й студент	50	48	40
4-й студент	49	47	35
5-й студент	50	44	40

У результаті виконання першої частини експерименту над простими формулами ми отримали такий результат: $\frac{50+49+50+49+50}{5} * 2 = 99,2\%$. По розрахункам

другої частини експерименту над формулами середньої складності отримано результат: $\frac{50+50+48+47+44}{5} * 2 = 95,6\%$.

Третя частина експерименту над складними формулами показала такий результат: $\frac{48+44+40+35+40}{5} * 2 = 82,8\%$.

Отже, можна зробити висновок, що розроблена програма озвучення математичних формул MathPlay показала хороший результат (92,5 %) для різних за складністю формул.

Учасники експерименту зазначили, що для них головною складністю було розуміння синтезованого голосу та швидкість відтворення, а також власне складність формули.

Здебільшого, причиною помилок при записі озвучених формул був синтезатор української мови «український голос UkrVox – Ігор». Відповідно, якщо буде розроблений якісний синтезатор української мови, який максимально наблизить звучання до

природного голосу, то відсоток розуміння озвученої формули з першого разу зросте до максимального. Проблема розроблення якісного синтезатора української мови залишається актуальною і не розв'язаною до кінця вже тривалий час.

Дослідження проблеми автоматичного озвучення математичних формул та символів показало, що із застосуванням системи правил перетворення тегів MathML розроблено якісну програму озвучення формул українською мовою (92,5 %) для різних за складністю формул.

Отримані результати дають змогу застосувати програму озвучення на практиці для навчання незрячих користувачів. Також, стало можливим використовувати озвучені українським голосом формули для створення навчальних підручників точних дисциплін у форматі DAISY засобами книг, що «розмовляють».

Подальше вдосконалення засобів озвучення математичних формул українською мовою вимагає розроблення якісного синтезатора української мови та вдосконалення методів підготовки даних.

ЛІТЕРАТУРА

1. INFTY: an integrated OCR system for mathematical documents. in Proceedings of the 2003 ACM symposium on Document engineering, (Grenoble, France, 2003), ACM Press, 95-104 / [Suzuki M., Tamari F., Fukuda R., Uchida S. and Kanahori T.]. – Режим доступу : http://www.inftyproject.org/articles/2003_DocEng_Suzuki.zip.
2. W3C. MathML Software [Електронний ресурс]. – Режим доступу : <http://www.w3.org/Math/ implementations.html>.
3. Ayars, J., Bulterman, D., Cohen, A., Day, K., Hodge, E., Hoschka, P., Huche, E., Jourdan, M., Kim, M., Kubota, K., Lanphier, R., Layaida, N., Michel, T., Newman, D., van Ossenbruggen, J., Rutledge, L., Saccocio, B., Schmitz, P. and ten Kate, W. Synchronized Multimedia Integration Language (SMIL 2.0) – [Second Edition], 2005. – Режим доступу : <http://www.w3.org/TR/2005/REC-SMIL2-20050107>.
4. Design Science. MathType. – Режим доступу : <http://www.dessci.com/mathtype>.
5. Design Science. WebEQ. – Режим доступу : <http://www.dessci.com/webeq>.
6. Brown, P. F., Pietra, S. A. D., Pietra, V. J. D., and Mercer, R. L. (1993). The mathematics of statistical machine translation. Computational Linguistics, 19(2):263–313.
7. Hutchins J. The evolution of machine translation systems / W. John Hutchins // Practical experience of machine translation, V. Lawson (ed.) North-Holland Publishing Company / © ASLIB, 1982.
8. MathML2 recommendation. Електронний ресурс. – Режим доступу : <http://www.w3.org/TR/MathML2>.
9. Катанова Т. Н. Набор и верстка математических текстов / Т. Н. Катанова, А. П. Шестаков // Информатика. – 2008. – № 12(565). – 48 с.
10. Ron Ausbrooks. Mathematical Markup Language (MathML) Version 2.0 (Second Edition) / Ron Ausbrooks, Stephen Buswell, David Carlisle et al. // W3C Recommendation 21 October 2003. – Режим доступу : <http://www.w3.org/TR/MathML2>.
11. OpenMath web site [Електронний ресурс]. – Режим доступу : <http://www.openmath.org/>.
12. Ian Hutchinson. Web Publishing Mathematics With MathML [Електронний ресурс] / Hutchinson Ian // IAP 2004. – Режим доступу : <http://web.m.it.edu/acs/iap05/mathml/mathmlfuture.pdf>.
13. Baker J., Sexton A. and V. Sorge «A linear grammar approach to mathematical formula recognition from PDF» in Proc. of Intelligent Computer Mathematics, 2009.
14. Josef B. Baker, Alan P. Sexton, Volker Sorge, Masakazu Suzuki : Comparing Approaches to Mathematical Document Analysis from PDF. ICDAR 2011: 463-467 [Електронний ресурс]. – Режим доступу : <http://www.inftyproject.org/en/index.html>.
15. [Електронний ресурс]. – Режим доступу : <https://www.dessci.com/en/products/mathdaisy/>
16. Kawamura Hiroshi DAISY: a better way to read, a better way to publish – a contribution of libraries serving persons with print disabilities/ Hiroshi Kawamura // Proceedings of the World Library and Information Congress: 72nd IFLA General Conference and Council. Meeting 91: Information Technology with Audiovisual and Multimedia and National Libraries (part 1), 20-24 August 2006 у, Korea, Seoul [Електронний ресурс]. – Режим доступу: <http://archive.ifla.org/IV/ifla72/papers/091-Kawamura-en.pdf>
17. Konev V. V. The Elements of Mathematics. Textbook. The Second Edition. Tomsk : TPU Press, 2001, 140 pp.
18. Електронний ресурс. – Режим доступу : <http://mytts.forum2x2.ru/t124-topic>.

Рецензенти: Кондратенко Ю. П., д.т.н., проф.;
Гожий О. П., к.т.н., доц.

© Давидов М. В.,
Лозицький О. А.,
Нікольський Ю. В., 2013

Дата надходження статті до редколегії 17.01.2013 р.

ДАВИДОВ М. В., докторант кафедри Інформаційних систем та мереж Національного університету «Львівська політехніка», м. Львів, Україна.

ЛОЗИЦЬКИЙ О. А., асистент кафедри Інформаційних систем та мереж Національного університету «Львівська політехніка», м. Львів, Україна.

НІКОЛЬСЬКИЙ Ю. В., д.т.н., професор кафедри Інформаційних систем та мереж Національного університету «Львівська політехніка», м. Львів, Україна.