

УДК 004.415.3; 544.1

Т. Нерода

Українська академія друкарства

МЕТОД МОДЕЛЮВАННЯ НОТАЦІЇ ФОРМАТУ АВТОМАТИЗОВАНОГО РЕДАКТОРА ХІМІЧНИХ ВИРАЗІВ

Розглянуто особливості структурних співвідношень між кодовими моделями інформаційних об'єктів автоматизованого редактора хімічних виразів для комп'ютерно-видавничих систем. Запропоновано математичні моделі позиціювання кодових записів, об'єднаних в окрему логічно організовану структуру формату даних прикладної програми хімічного редактора. Описано метод моделювання основних компонентів нотації формату.

Умовні позначення

ЗАГ – записи тега ЗАГОЛОВОК;

Іав – поле ідентифікатора тега АЦИКЛІЧНІ ВИРАЗИ;

Зк – запис координат поточного виразу;

n – кількість полів ациклічних виразів у тегу, покази локального лічильника;

ℓ – кількість вершин поточного циклу з ациклічним входженням;

m – кількість моделей циклу у формульному об'єкті, покази локального лічильника;

k – кількість полів незалежних графічних об'єктів у тегу, покази локального лічильника;

ЛчАЦ – поле лічильника ациклічних об'єктів;

ЗКМВ – запис кодової моделі поточного виразу;

ЗПЦ – записи параметрів циклу;

ЛчЦФ – поле лічильника моноциклічних фрагментів;

ЗКМА – запис кодової моделі ациклічного входження вершини циклу;

ЗПО – запис параметрів незалежного об'єкта;
ЗКМО – запис кодової моделі поточного об'єкта;
ЛчНГО – поле лічильника об'єктів;
ВХВ – висота виразу.

Вступ

Однією з рис розвитку ринкової економіки є поширення малих друкарських систем, які мають справу з широким асортиментом поліграфічної продукції (підручниками, довідниками, інформаційними виданнями, науковою літературою тощо), що містить значний відсоток нелінійних текстів. Скорочення термінів випуску і собівартості такої продукції можливе лише за умови правильного використання наявних та створення нових автоматизованих систем підготовки видань. Отже, постає актуальна проблема моделювання інформаційних компонентів програм комп'ютерно-видавничих систем (КВС), що ґрунтуватимуться на традиціях національного книгодрукування, вираженому естетичному оформленні, алгебрі логіки і граматики формальних мов [1].

Обґрунтування необхідності побудови формату даних

При моделюванні зображення структурних формул в середовищі автоматизованого редактора хімічних виразів постає проблема збереження формульних об'єктів для наступної їх візуалізації, модифікації та впровадження в документ комп'ютерно-видавничих систем. При такому збереженні інформація про сукупні параметри ФО повинна подаватись у пам'яті ЕОМ незалежними закритими блоками кодів, організованими в індивідуальну композицію даних, названу форматом прикладної програми. Як елемент специфікації формульного об'єкта, що визначатиме його тип, цей формат запропоновано реєструвати під ім'ям *.arb*.

Щоб забезпечити збереження різних типів даних – складно-структурованого тексту, хімічно орієнтованої графіки тощо, формат редактора хімічних виразів повинен ґрунтуватися на синтаксисі систем кодування зображень хімічних формул [2] та оперувати потоками даних, прийнятими у використовуваній операційній системі. Формат *.arb* складається з послідовності даних, названих тегами. Структура, розмір та позиція тегів інформаційних об'єктів хімічного редактора неоднакові і змінюються для різних розокремлених виразів. Розташування тега задається абсолютним зміщенням від ідентифікатора формату. Теги містять набори зв'язаних полів, складених з записів, які формують цілісну характеристику однорідних інформаційних об'єктів.

Як і для більшості засобів збереження комп'ютерних даних, теги формату хімічного редактора впорядковано у певну базову структуру: заголовок, розділ даних, маркер кінцівки. Наведена структура дозволяє коректно зберігати дані автоматизованого редактора і належним чином інтерпретувати їх при візуалізації.

Моделювання структурних компонентів нотації формату

Таким чином, нотацію формату *.arb* структуровано відповідно до якісних ознак, що складають специфіку автоматизованого редактора хімічних виразів.

Після завантаження прикладної програми передусім налаштовуються поліграфічні параметри моделей візуалізації інформаційних об'єктів, відбувається присвоєння їх аргументам відповідних значень. При цьому в засобі збереження даних середовища прикладної програми формується тег ЗАГОЛОВОК і створюється більшість його записів. Далі прикладна програма переходить в режим очікування дій набірника. В разі активізації шаблону піктографічного меню [3] запускається загальний лічильник інформаційних об'єктів хімічного редактора та відбувається ідентифікація введеного об'єкта.

Якщо розпочато синтез моделі ациклічного виразу, ініціалізується процедура опрацювання компонентів ациклічного ланцюга – текстових блоків, типу, напрямку і порядку зв'язків, наявності й особливості галуження. Після цього формується заголовковий ідентифікатор тега даних ациклічного виразу. До тега долучаються записи відно-

сних координат та кодової моделі вуглецевого скелету поточного виразу. Цей код представлено лінійним записом, який зберігає структуру вуглецевого скелету головного ланцюга зображення виразу із врахуванням розгалужень. Такий метод запису полягає у поданні ідентифікаторів хімічних зв'язків, обраних з піктографічного меню, як операторів, що сполучають аргументи – символічні позначення хімічних елементів та їх якісного і кількісного складу, знаків математичних дій тощо, об'єднані у текстові блоки [4]. Код ациклічного виразу охоплюється ідентифікаторами його початку і кінця. Процедура моделювання поля виразу закінчується інкрементом лічильника інформаційних об'єктів тега АЦИКЛІЧНІ ВИРАЗИ.

В разі опрацювання циклу відбувається ідентифікація кількості вершин у полі тега ЦИКЛІЧНІ ФРАГМЕНТИ; для поточної вершини аналізується наявність, тип та порядок зв'язку. Ці дані формуються у кодову модель циклу і долучаються до записів параметрів циклу; у кодовій моделі циклу передбачено зберігати характеристики кільця, до яких відноситься розташування вершин базового многокутника і порядок та напрямок кожного зокрема зв'язку сполуки. Згідно з розробленою системою кодування, такі характеристики зберігаються у масиві: на перетині рядка та стовпця під головною діагоналлю матриці записується порядок зв'язку між відповідними вершинами; особливості зображення стереозв'язку у циклі Хеурса, а саме тип та напрямок, зазначаються у частині матриці, розташованій над головною діагоналлю. При цьому відносна ширина знака моноциклу фіксується у відповідній позиції кодового запису ациклічного виразу; покази лічильника циклів збільшуються на одиницю. Якщо для поточного циклу моделюється символічний блок чи замісник, це зазначається кодом ациклічного входження в теґовому полі запису відповідної вершини цього циклу.

Модифікації теґа НЕЗАЛЕЖНІ ГРАФІЧНІ ОБ'ЄКТИ передують моделювання об'єкта за вказаними точками та долучення їх атрибутів до запису параметрів цього об'єкта. Особливості наступної візуалізації зазначаються у кодовому записі, що разом із записом параметрів складає поле поточного графічного об'єкта, число яких зберігається у лічильнику інформаційних об'єктів розглядуваного теґа.

Створення наступних об'єктів чи модифікація наявних відбувається за описаним алгоритмом (рис. 1) шляхом синтезу відповідних кодових моделей та долучення їх до полів інформаційної структури наявної нотації формату.

При вивантаженні прикладна програма передусім перевіряє стан загального лічильника об'єктів, покази якого повинні збігатися з показами індивідуальних лічильників теґів даних. В разі невідповідності середовищем автоматизованого редактора видається повідомлення і з нотації формату вилучається поле записів об'єкта, який призвів до збою.

Далі передбачено розрахунок позицій теґів даних. Теґ АЦИКЛІЧНІ ВИРАЗИ розпочинається полем, дев'ятим від ідентифікатора формату. Позиція поля ідентифікатора наступного теґа розраховується таким чином:

$$T_{цф} = \sum (ЗАГ) + I_{AB} + ЛчАЦ + \sum_{i=1}^n (3k_i + 3КМВ_i). \quad (1)$$

Поле ідентифікатора теґа НЕЗАЛЕЖНІ ГРАФІЧНІ ОБ'ЄКТИ розташоване у позиції

$$T_{НГО} = T_{цф} + ЛчЦФ + \sum_{i=1}^m \left(3ПЦ_i + \sum_{j=1}^l 3КМА_{ij} \right). \quad (2)$$

Відтак, маркер завершення потоку даних, запис якого розпочинає теґ КІНЦІВКА, зміщений від ідентифікатора формату на величину

$$T_{EOF} = T_{НГО} + ЛчНГО + \sum_{i=1}^k (3КМО_i + 3ПО_i). \quad (3)$$

Після маркера EOF теґ КІНЦІВКА містить декілька записів доповняльних відомостей, які не враховуються при візуалізації.

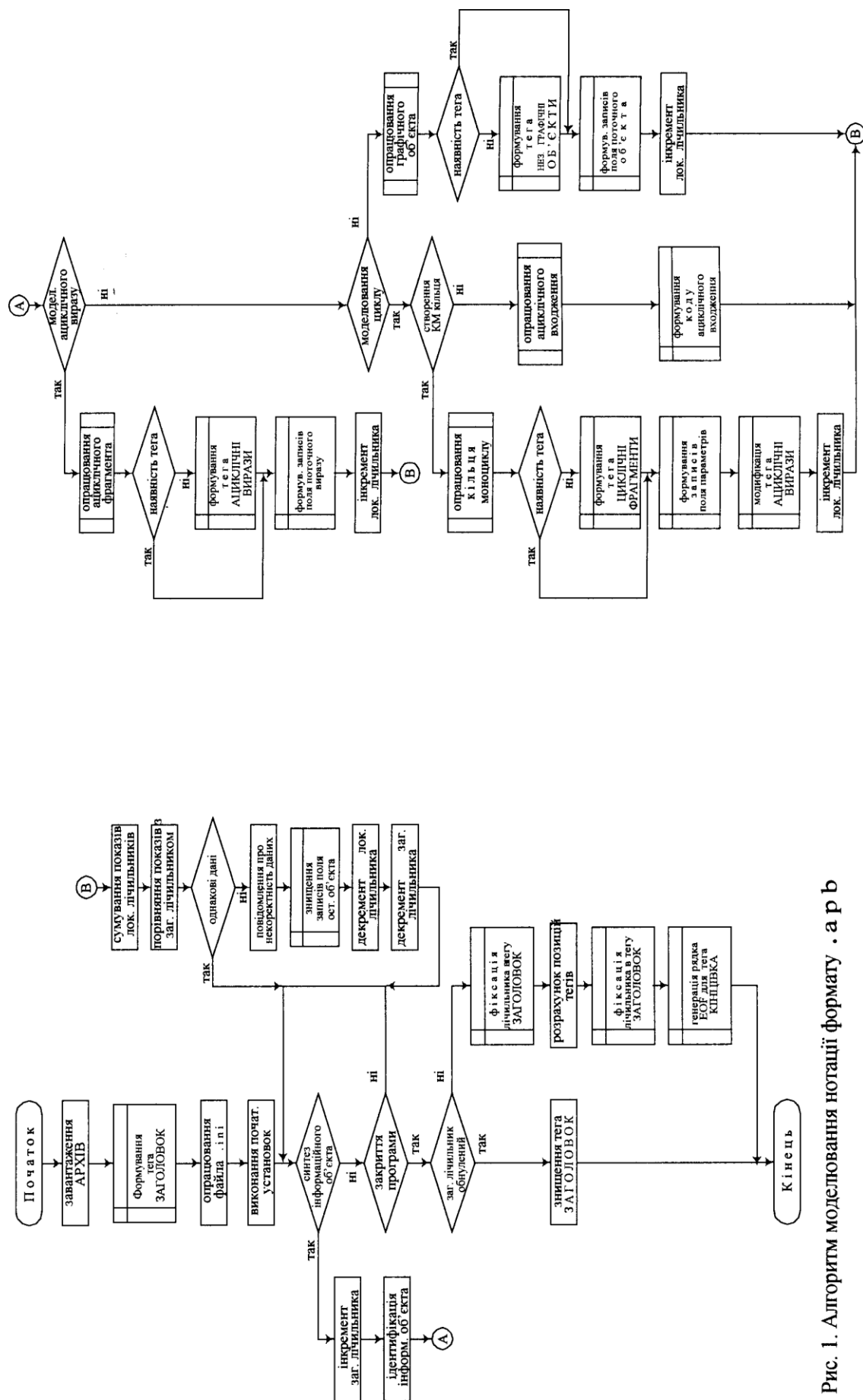


Рис. 1. Алгоритм моделювання нотатції формату . a r b

Рис. 1. Алгоритм моделювання нотатції формату . a r b

Розраховані таким чином записи позицій тегів нотації формату . a r b зазначаються у відповідних записах тега ЗАГОЛОВОК і надалі вважатимуться основними засобами доступу до окремих параметрів моделей інформаційних складових формульного об'єкта.

Візуалізація формульного об'єкта

Загалом, основна проблема наступного розшифрування формату полягає в існуванні декількох рівнів ієрархії. При зчитуванні формату редактор хімічних виразів віднаходить потрібні дані, використовуючи прийняті в специфікації угоди і дефініції окремих компонентів. Також слід зазначити, що записи формату . a r b обмежені символами повернення каретки, щоб їх зручно було переносити до оперативного списку активної прикладної програми.

У модулі читання формату середовища автоматизованого редактора хімічних виразів передбачено ряд наведених далі функцій, що можуть виявити помилки коду та їх причини і запобігти зависанню прикладної програми:

- ідентифікація формату виключно за записами тега ЗАГОЛОВОК;
- ідентифікація кінця нотації за маркером EOF тега КІНЦІВКА;
- перевірка записів розташування тегів на відповідність значенням позицій;
- перевірка значення загального лічильника об'єктів на відповідність сумарному значенню показів локальних лічильників.

Якщо деякий запис некоректний, модуль читання спробує проаналізувати решту нотації і в разі негативного результату видасть відповідне повідомлення. В разі позитивного результату модуль записування формату відповідним чином відкоригує нотацію. Далі розпочнеться процедура візуалізації формульного об'єкта відповідно до початкових установок: модуль читання формату ідентифікує підзаголовок тега АЦИКЛІЧНІ ВИРАЗИ і опрацює значення лічильника об'єктів, запис якого розташований в наступному полі. Відповідно до показів локального лічильника програма сканує поле поточного об'єкта. Першим записом цього поля є відносна позиція виразу, зміщена за рахунок наявності графічних моделей зображень попередніх виразів. Так, координати n -го об'єкта з врахуванням початкових установок поліграфічних даних розраховується таким чином:

$$\begin{aligned} x_n &= const, \\ y_n &= \sum_{i=1}^{n-1} BXB_i. \end{aligned} \quad (4)$$

Після розрахунку позиції та аналізу кодової моделі поточного хімічного виразу відбувається власне моделювання зображень його компонентів. Побудова графічної моделі завершується декрементом локального лічильника. Його обнулення свідчить про завершення процедури візуалізації об'єктів даного тега та ідентифікацію наступного за формулою (1). Тут також передусім опрацюється поле локального лічильника, що його покази визначатимуть поточний моноцикл, активізується процедура аналізу параметрів кодової моделі циклу відповідно до початкових налаштувань, результатом чого є побудова моделі візуалізації, – визначення сукупності команд, які за методами формування зображення моноциклу керують створенням графічного образу поточного об'єкта. За наявності ациклічного входження підпрограми автоматизованого редактора опрацюють його код та моделюють графічний образ.

Після обнулення лічильника тега циклів модуль читання позиціюється у підзаголовку поля НЕЗАЛЕЖНІ ГРАФІЧНІ ОБ'ЄКТИ (2). Для кожного об'єкта тут також відбувається аналіз коду і розрахунок параметрів. За результатами будується модель візуалізації, яка відповідно до кодової моделі формує графічний образ зображення об'єкта. Обнулення локального лічильника свідчить про відсутність інформаційних об'єктів для візуалізації у даному тегу та у форматі в цілому.

Висновки

Отже, моделювання нотації розглянутого формату за розробленим методом відбувається програмним чином, використовуючи мінімальний набір даних, придатних до опрацювання. Результатом процедури моделювання нотації формульного об'єкта є індивідуальна логічно організована структура даних у пам'яті ЕОМ, яка здатна підтримувати в документі КВС двомірну графічну модель вихідного зображення хімічного виразу, виконану з суворим дотриманням вітчизняних технологічних вимог правильного поліграфічного відтворення друкованої продукції [1], ретельно закладених у програмне середовище автоматизованого редактора хімічних виразів.

Ця графічна модель, як зв'язаний з програмою-сервером та впроваджений в програму верстання об'єкт, придатна до модифікації окремих її компонентів разом з коригуванням відповідних записів у нотації формату. Рациональність розробленого формату, зокрема передбачення механізму виявлення та усунення помилок, є вирішальним чинником, що визначає швидкість, з якою логічний блок даних автоматизованого редактора хімічних виразів, розташований у файлі публікації КВС, може бути ідентифікований і прочитаний, розмір зайнятого ним дискового простору і простоту доступу через формульний об'єкт до середовища програми-сервера за даними цього блоку.

Таким чином, запропоновану структуру нотації, змодельовану з використанням розроблених методів кодування зображень хімічних сполук [4], побудовані математичні моделі та алгоритм формування нотації доцільно застосувати при створенні процедурних модулів запису й читання формату . a r b автоматизованого редактора хімічних виразів для комп'ютерно-видавничих систем.

The features of structural relations between code models of information objects of the automatized editor of chemical expressions for computer publishing systems surveyed. The mathematical models of positioning of code records, integrated in the separate logically organized structure of data format of an application of the chemical editor are offered. The method of simulation of main components of the notation of the format is explained.

Література

1. Партико З.В. Загальне редагування: нормативні основи. – Львів: Афіша, 2001. – 416 с.
2. Нерода Т. Аналіз та класифікація систем кодування хімічних формул // Наукові записки УАД. – Львів, 2002. – №5. – С. 107-110.
3. Нерода Т. Елементи структурної організації редактора хімічних виразів // Комп'ютерні технології друкарства. – Львів, 2002. – №7. – С. 116-121.
4. Нерода Т. Методи кодування зображень хімічних виразів для комп'ютерно-видавничих систем // Комп'ютерні технології друкарства. – Львів, 2002. – №8. – С. 119-124.

Одержано 22.01.04 р.