

УДК 004.921

**Б. В. Дурняк**

*Українська академія друкарства*

**О. М. Назаренко**

*Кримський інститут інформаційно-поліграфічних технологій УАД*

**РОЗРОБКА АЛГОРИТМІВ СИНТЕЗУ ІНФОРМАЦІЙНИХ  
КОМПОНЕНТ З ІНШИМИ КОМПОНЕНТАМИ  
В РАМКАХ МОДЕЛІ ЯКОСТІ КНИГИ**

*Розроблено алгоритм синтезу інформаційних компонент моделі якості книги з іншими компонентами, представлено послідовність реалізації правил текстових перетворень при семантичному аналізі інформаційної моделі якості, наведено функціональну схему реалізації текстових перетворень.*

***Системний аналіз, інформаційні моделі, інтерпретаційні описи***

Синтез інформаційних засобів, що об'єднуються в інформаційну модель ІМІ, реалізується відповідно до функціональної орієнтації окремих інформаційних компонент, які являють собою семантичний словник (Sc) і системи: правил перетворень текстових, інтерпретаційних описів ( $\Sigma$ ); правил виводу текстових описів ( $\xi$ ); семантичного аналізу текстових описів; інформаційного опису компонент моделі якості книги МІК.

Семантичний словник — це впорядкована множина текстових описів у нормалізованій формі окремих елементів МІК та системи, що описує відповідну предметну область [1]. Предметну область  $W_i$ , крім моделі МІК, складають вхідні дані, які використовуються моделями-компонентами, наприклад GМІ, LМІ, та ін. компонентами, що розширюють і доповнюють МІК.

Система  $\Sigma$  відображає особливості застосування текстових описів при розв'язуванні задач, що використовують МІК як базовий засіб розв'язку. Правила перетворень являють собою перестановку, додавання та елімінацію фраз. Коректність їх використання ґрунтується на аналізі семантичних параметрів, що полягає в перевірці допустимості значень відповідних параметрів, які можуть змінюватися в результаті відповідних перетворень.

Система перетворень описів інтерпретації окремих компонент  $W$  ґрунтується на використанні спеціальних схем, які відображають значення семантичних параметрів, котрим повинні відповідати фрази, що відповідно до окремо введеного перетворення можуть доповнювати фрагмент тексту, яким передбачається розширяти фрагмент, що піддається перетворенню, та граматичні типи слів, з яких повинен формуватися відповідний фрагмент. Самі слова у відповідних схемах не подаються, а повинні вибиратися із семантичних

словників, використовуваних для опису предметної області. Формально таке правило виводу текстового розширення для деякого фрагмента  $(\varphi_1 * \dots * \varphi_k)$  можна подати у вигляді співвідношення

$$\left\{ (\varphi_1 * \dots * \varphi_k) \& [\sigma_i(\varphi_1 * \dots * \varphi_k), \dots, \sigma_n(\varphi_1 * \dots * \varphi_k)] \& [\gamma(\varphi_1 * \dots * \varphi_k)] \right\} \rightarrow (1) \\ \rightarrow [(\varphi_1 * \dots * \varphi_k) * (\varphi_1^* * \dots * \varphi_m^*)],$$

де  $\sigma_i(\varphi_1 * \dots * \varphi_k)$  — значення семантичного параметра  $\sigma_i$  фрази;  $(\varphi_1 * \dots * \varphi_k)$ ,  $\gamma(\varphi_1 * \dots * \varphi_k)$  — граматичні типи ключових слів, які повинні використовуватись у фрагменті, що передбачає розширяти певну фразу або деякий фрагмент тексту  $(\varphi_1 * \dots * \varphi_k)$ .

Це правило визначає спосіб вибору нового фрагмента, який потрібно побудувати при реалізації операції розширення текстових описів. Рівень нормалізації правил виводу в математичній логіці нових логічних формул має такий рівень абстракції, що не дозволяє висловлюватися про правила вибору необхідних елементів у процесі побудови або реалізації процедури виводу [3]. Тому в математичній логіці процедури вибору чергових елементів для побудови ланцюга виводу являють собою процеси перебору. У рамках системи виводу, використовуваної для перетворень текстових описів інтерпретації  $j(x_i)$ , такі процедури вибору чергових елементів для поточного кроку побудови процесу виводу можуть визначатися певними правилами та вимогами. У даному випадку відповідні правила та вимоги не можуть існувати порівняно незалежно від самих правил перетворень типу (1). Відповідні вимоги синтезовані з цими правилами і, по суті, відображають механізм реалізації таких правил. Тому на прикладі правила перетворення (1) розглянемо спосіб реалізації певного правила, який регламентується відповідними вимогами та додатковими правилами. Послідовність реалізації правила перетворень може бути подана таким чином:

1. Вибираються деяка кількість речень  $\Psi_i$ , одне речення або одна фраза  $\varphi_i$ , які необхідно розширити.

2. Визначаються значення семантичних параметрів для всіх фраз, що складають абзац тексту, до якого передбачається реалізувати текстове розширення.

3. На основі визначених значень семантичних параметрів обчислюються динамічні величини зміни цих параметрів у тексті, який аналізується і передбачається доповнювати, або параметри  $\rho\sigma_i$ .

4. Визначаються граматичні схеми кожної з фраз, що складають текст, який аналізується  $\gamma_i(\varphi_i)$ .

5. З визначених граматичних схем обчислюється міра поширеності окремих типів слів та граматичних схем, використовуваних у фразах тексту, що аналізується.

6. З ознак, які описують граматичні схеми, реалізується вибір чергових слів ідентифікаторів із словника  $S_c$  з функціональної групи, до якої відноситься найпоширеніший граматичний тип слова.

7. На основі підбраного ключового слова для фрази, яку передбачається формувати, зі словника  $S_c$  вибираються слова відповідно до ієрархії їх значимості у фразі, що вибирається на основі величини її пріоритету, встановленого в пункті 2 даного алгоритму, і ця процедура повторюється доти, доки згідно з вибраною граматичною схемою фрази прототипу не буде сформована відповідна фраза.

8. Після завершення процесу формування чергової фрази здійснюється перевірка нівелювання параметрів причини, що обумовила потребу у формуванні розширеного відповідного фрагмента тексту.

9. Коли параметри, які визначають необхідність розширення текстового опису фрагмента  $j(x_i)$ , нівельовані, то відповідний процес побудови нових фраз завершується. Якщо відповідні параметри залишаються значимими, то алгоритм переходить до аналізу фрази, яка має наступний або такий самий пріоритет чи значимість, і робота алгоритму повторюється.

10. Якщо параметри причини, що обумовлюють потребу в розширенні фрагмента, у достатній мірі не вдалося зменшити, то алгоритм передає управління системі, яка визначає наступний спосіб модифікації відповідного фрагмента абзацу тексту.

Розглянемо і чіткіше визначимо причини, які обумовлюють необхідність модифікації текстових фрагментів, що застосовуються в МЖК і являють собою процеси узгодження інтерпретаційних описів компонент, використовуваних у моделі.

Першою причиною, що зумовлює необхідність інтерпретаційного узгодження двох процесів або двох компонент, є вихід параметрів, якими обмінюються два фактори в МЖК, за допустимі, визначені границі їх значень. У цьому випадку параметрами, що спричиняють модифікацію  $j(x_i)$ , є величина кількості недопустимих відхилень значень відповідних параметрів у рамках компонент, що передає відповідний параметр іншій компоненті МЖК. Наступним параметром, який характеризує потребу в реалізації процесів інтерпретаційного узгодження, є частота виникнення відповідних відхилень у значеннях параметрів, що передаються від однієї компоненти до іншої в процесі функціонування МЖК, та частота виходу за допустимі значення поточного параметра, котрий у даний момент вийшов за допустимі границі. Позначимо, відповідно, ці параметри  $\Delta P_i^j$ ,  $\delta P_i$ . Очевидно, що третім параметром такої причини є величина відхилення значення параметра за встановлені границі допустимих значень, яку позначимо  $\Delta P_i^j$ . Для того щоб можна було ініціювати відповідне розширення, достатньо, аби актуальним виявився один параметр  $\Delta P_i^j$ , де  $i$  — номер параметра або його ідентифікатор. Виходячи з наведених прикладів, можна стверджувати, що ці параметри визначають у сукупності значимість, або величину важливості відповідної причини. Такий параметр, як величина значимості причини, що викликає проведення перетворень з узгодження інтерпретаційних описів інтерфейсів двох компонент МЖК, є дуже важливим, оскільки вказана значимість визначає величину текстового опису, який повинен аналізуватися

при проведенні перетворень  $j(x_i)$ , пов'язаних з реалізацією інтерпретаційних узгоджень двох окремих компонент. Відомо, що мінімальною частиною  $j(x_i)$  є одна фраза. Збільшення розміру  $j(x_i)$  полягає у виділенні такої одиниці, як одне речення  $\Psi_i(\varphi_{i1} * \dots * \varphi_{ik})$ . Наступною одиницею такого збільшення є один абзац, який складається з двох або більше речень  $\Phi(\psi_{i1}, \dots, \psi_{in})$ . Якщо значимість причини, яку позначатимемо символом  $D^i$  ( $i$  — ідентифікатор причини), є такою, що однієї фрази для аналізу замало, що зумовлює реалізацію відповідного розширення, то наступною компонентою, котра вважається базовою для аналізу, на основі якого проводиться розширення, є речення  $\psi_i$ . Очевидно, що випадок, коли одна фраза складає одне речення, або  $\varphi_i = \psi_i$ , не розглядається. Це означає, що збільшення кількості фраз, які вибираються для аналізу в процесі реалізації розширення, не є рівномірним, а від однієї фрази здійснюється перехід до одного речення. Подібна ситуація має місце при переході від одного речення до одного абзацу. Таким чином, можна записати, що збільшення розміру тексту для проведення аналізу в процесі його модифікації реалізується згідно з наступною схемою:

$$\varphi_i \rightarrow [\Psi_s(\varphi_{i1}, \dots, \varphi_{in}) \& (n > 1)] \rightarrow [\Phi_i(\psi_{i1}, \dots, \psi_{in}) \& (n > 1)].$$

Оскільки значимість причини є суттєвим фактором для інтерпретаційного розширення, то сформуємо співвідношення для визначення  $D^v$ , де  $v$  означає причину, що полягає у виході значень параметрів за допустимі границі:

$$D^v = f(\Delta P, \delta P_p, \Delta p_i). \quad (2)$$

Пороги величини значимості  $D^v$ , за якими можна приймати рішення щодо розміру фрагменту тексту, що вибирається як посилення для співвідношення (1), визначаються на основі експериментальних даних. Функція  $f$  у співвідношенні (2) вибирається таким чином, щоб найбільший вплив на визначення величини  $D^v$  мав параметр  $\Delta P$ , наступне — за величиною значення параметра  $\delta p_i$ . Враховуючи це, можна записати співвідношення

$$D^v = f[(\alpha \Delta P), (\alpha \beta^{-1} \delta P_i), \Delta p_i],$$

де коефіцієнти  $\alpha$  і  $\beta$ , що визначають ваги відповідних параметрів, є цілими числами і знаходяться у співвідношенні  $\alpha > \beta$ .

Другою причиною необхідності модифікації  $j(x_i)$  є виникнення дисбалансу в GMJ, який визначається як збільшення розбалансованості GMJ за деякий заданий поріг. Цю причину позначатимемо символом  $D^R$ ; параметрами, що визначають величину дисбалансу MJK, є:

величина дисбалансу, яка встановлюється як різниця між величиною поточного значення  $d^B$  і порогом допустимого значення розбалансування  $\Delta d^B$ , що формально можна записати співвідношенням  $dR = d^B - \Delta d^B$ ;

тип класу причин, котрі призводять до недопустимого розбалансування, якими можуть бути технологічні процеси та їх параметри, параметри конструкції книги, що позначимо  $d^L$ ;

міра локальності причини виникнення дисбалансу, яку позначатимемо  $d^L$ .

Визначення типу причини дисбалансу має велике практичне значення, оскільки різними методами треба вибирати не відповідні параметри, а з того чи іншого параметра той або інший клас текстових описів.

Міра локальності дисбалансу  $d^L$  — класи походження причин його виникнення з точки зору типів параметрів, які визначають відповідний дисбаланс. Відомо, що параметри, які можуть впливати на якість книги, мусять бути тією чи іншою мірою взаємозв'язаними. Тому міра локальності визначає, яка кількість параметрів впливає на параметр  $J(PV_i)$  та як вони взаємозв'язані між собою. Цей параметр досить важливий не тільки для визначення величини фрагмента тексту посилання, а й способу протидії чи елімінації такого відхилення, як дисбаланс, що дозволяє сформувати додаткові умови для вибору елементів розширення текстового фрагмента, котре формується в рамках співвідношення (1).

Наступною причиною ініціації процесів модифікації текстових описів є виникнення недопустимих відхилень у значеннях семантичних параметрів текстових описів, що уже використовуються в рамках ІМЖ. Очевидно, що ця причина  $D^6$  може виникнути тільки тоді, коли в процесі функціонування МЖК реалізується семантичний аналіз текстових описів. Отож розглянемо детально такі випадки. Семантичний аналіз текстових описів  $j(x_i)$  проводиться незалежно від типу моделі, яка ініціюється в поточний момент, коли:

у результаті функціонування деякого фрагмента моделі остання переходить у режим реєстрації аномальної ситуації;

використовується LMJ і в процесі функціонування виникає логічна суперечність;

у рамках функціонування GMJ виявляється, що для деякого вузла відсутнє алгоритмічне наповнення;

у процесі функціонування окремого фрагмента аналітичної моделі АМЖ виявиться, що остання при заданих початкових умовах не має розв'язку;

МЖК у процесі функціонування передає управління системі аналізу текстових описів інтерпретації окремих фрагментів моделі, які не піддаються формальному опису, оскільки рівень абстракції такого опису дискредитує відповідний фрагмент.

Довільна модель функціонує шляхом реалізації певного алгоритму [2]. Коли йдеться про алгоритми, які повинні проводити аналіз досить складних об'єктів, то завжди може виникнути ситуація, вирішення якої алгоритму виявиться не під силу, через те що необхідних даних для його функціонування на момент проектування немає. У цьому випадку алгоритм реєструє факт неможливості продовжувати процес функціонування, що проявляється ідентифікацією виникнення аномального стану алгоритму. Ця інформація передається в ІМЖ. Модель ІМЖ вміщає текстові описи функціонування всіх алгоритмів, що реалізуються в МЖК, та описи всіх даних, які існують у відповідному фрагменті предметної області. При цьому ініціюється процес аналізу відповідного текстового опису  $\Phi_i(\Psi_1, \dots, \Psi_n)$  для виявлення необхідних додаткових даних,

завдяки яким стало б можливим уникнути виникнення аномальної ситуації в алгоритмі  $A_i$ .

При використанні моделі LMJ, яка являє собою систему логічних співвідношень, що описують різні компоненти предметної області  $W$  шляхом її інтерполяції логічними співвідношеннями, логічна суперечність може виникати тільки за рахунок неповної або неточної інтерпретації, на основі якої формуються аксіоми для відповідного фрагмента моделі МJK. Завдяки аналізу текстових описів, які найбільш адекватні певній предметній області, оскільки останні можуть поповнюватися за рахунок нової інформації, котру користувач постійно вводить у МJK, стає можливим усунути відповідну суперечність шляхом модифікації аксіом на основі нових даних про об'єкт досліджень.

Формально такий процес у логічній формі можна записати у вигляді співвідношень

$$\begin{aligned} & \{[L_i(x_{i1}, \dots, x_{ik}) \& [J(x_{i1}), \dots, J(x_{in})]]\} \rightarrow M[J(x_{i1}, \dots, x_{im})] \rightarrow \\ & \rightarrow \{[L_i(x_{i1}, \dots, x_{ik}) \& L_i^*(x_{i1}, \dots, x_{ik})]\}, \end{aligned}$$

де  $L_i^*(x_{i1}, \dots, x_{ik})$  — логічна аксіома, модифікована на основі даних, які призвели до зміни інтерпретації окремих елементів у формулі  $L_i(x_{i1}, \dots, x_{ik})$ , завдяки чому можливість виникнення логічної суперечності елімінована у відповідному фрагменті LMJ.

Модель GMJ являє собою схему, яка об'єднує всі компоненти моделі МJK, необхідні для визначення параметра  $j(PVi)$ . Кожна з компонент відповідає окремому вузлу графового зображення. Якщо при розв'язанні задачі визначення  $j(PVi)$  вибрано шлях у певному графі, то для досягнення  $j(PVi)$  усі вузли, через які проходить цей шлях, повинні вміщати відповідний алгоритмічний засіб, що здійснює перетворення, котре формально описується співвідношенням

$$Pi = [(P_k^j \alpha_{ki}^j), \dots, (P_{k+i, i})], \quad (3)$$

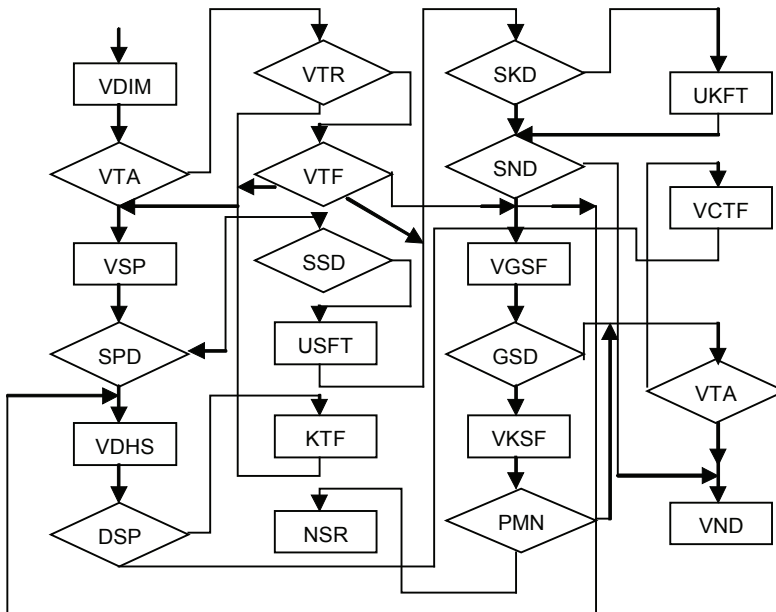
де індекс « $i$ » — вихідний параметр рівня ієрархії « $i$ »;  $k$  — рівень ієрархії, з якого виходить вектор, що входить у відповідний вузол (при цьому  $k < i$ );  $j$  — рівень ієрархії параметрів, що виходять з вузла. Таким чином, якщо  $t_i$  із співвідношення (3) виявився порожнім, то на основі даних про вхідні і вихідні параметри вузла вибирається потреба у фрагменті текстових описів, що знаходяться в  $[M]$ .

Неточність і невідповідність формалізованої частини МJK поточному стану предметної області є допустимим, оскільки формальні засоби формуються чи модифікуються тільки тоді, коли виявляються непридатними розв'язувати відповідну задачу. Переважно в рамках МJK аналізується можливість виникнення аномалії в процесі розв'язання окремої задачі на етапі передачі результатів, отриманих в одній задачі, на вхід наступної, яка відповідно до алгоритмів функціонування МJK повинна активізуватися. Коли йдеться про неузгодженість числових даних, то цей випадок уже розглядався. Але в рамках відповідного інтерфейсу зв'язку можуть передаватися не тільки числові величини, а й ті чи інші розпорядження, які можуть фізично представляти числа, а



інтерпретуватися з боку приймача як команда і т.д. Якщо відповідна команда немає необхідної інтерпретації з боку приймача, то вона може перейти в некерований стан. З огляду на це в таких випадках МЖК передає управління ІМЖ. Остання на основі отриманих від МЖК усіх необхідних даних передає управління РІЖК, де передусім вибирається текстовий фрагмент для проведення над ним відповідних перетворень й інших дій, пов'язаних з розв'язанням задачі щодо визначення величини  $j(PVi)$ .

Коли функціонує окрема модель АМЖ, яка не в стані розв'язати відповідну задачу, то це може обумовлюватися тим, що вхідні дані не відповідають вимогам, що є обов'язковими для відповідної АМЖ. Така нездатність може проявлятися в тому, що АМЖ переходить в аномальний стан, який передбачається алгоритмами, функціонуючими в АМЖ, або тим, що на виході АМЖ формуються дані, котрі не мають визначеної для АМЖ інтерпретації. Тому в таких випадках АМЖ передає управління ІМЖ через системи засобу МЖК чи безпосередньо. Оскільки прийнято вважати, що  $J(Wi)$  є найповнішим описом  $Wi$ , то ІМЖ у другому випадку формує текстовий опис інтерпретації, якої на поточний момент не вистачає, а в першому — на основі аналізу текстових описів реалізує текстові рекомендації з розширення аналітичних можливостей відповідної моделі АМЖ.



Функціональна блок-схема текстових перетворень, що реалізується в ІМЖ: VDSM — аналіз вхідних даних та активізація ІМЖ; VTA — визначення розміру тексту, який необхідно аналізувати; VTR — визначення, чи вибраний текст є реченням; VSP — визначення семантичних параметрів; VTF — визначення, чи вибраний текст є фразою; SPD — перевірка, чи семантичні параметри допусти-

мі; VDHS — визначення динамічних характеристик семантичних параметрів; USFT — усунення семантичних суперечностей у фрагменті тексту; SKD — визначення, чи семантичний конфлікт допустимий; VND — вихід з некоректних даних; UKFT — усунення конфлікту у фрагменті тексту; DSP — перевірка, чи динамічні семантичні параметри допустимі; KTF — корекція тексту фрагменту по динамічних семантичних параметрах; VGSF — визначення граматичних схем фраз; GSD — визначення, чи граматичні схеми допустимі; VKSF — вибір ключових слів для фрази; PMN — перевірка, чи причина необхідності модифікації нівельована; NSR — нормальне завершення розширення; VCTF — вибір чергового типу фраз; VTA — перевірка, чи всі типи фраз проаналізовані; SSD — визначення, чи семантична суперечність допустима; SND — визначення, чи семантична надмірність допустима.

1. Афанасьєва О. Ю. Методи семантичних перетворень в стеганосистемах / О. Ю. Афанасьєва // Моделювання та інформаційні технології : зб. наук. пр. (ПІМЕ. НАН України). — К., 2010. — Вип. 56. 2. Мишкіс А. Д. Элементы теории математических моделей / А. Д. Мишкіс. — М. : Наука, 1994. 3. Такедти Г. Теория доказательств / Г. Такедти. — М. : Мир, 1978.

## **РАЗРАБОТКА АЛГОРИТМОВ СИНТЕЗА ІНФОРМАЦІЙНИХ КОМПОНЕНТ С ДРУГИМИ КОМПОНЕНТАМИ В РАМКАХ МОДЕЛІ КАЧЕСТВА КНИГИ**

*Разработан алгоритм синтеза информационных компонент модели качества книги с другими компонентами, представлена последовательность реализации правил текстовых преобразований при семантическом анализе информационной модели качества, приведена функциональная схема реализации текстовых преобразований.*

## **DEVELOPMENT OF ALGORITHMS OF SYNTHESIS OF INFORMATIVE IS COMPONENT WITH OTHER COMPONENTS WITHIN THE FRAMEWORK OF MODEL OF QUALITY OF BOOK**

*The algorithm of synthesis of informative is developed component of model of quality of book with other components, the sequence of realization of rules of text transformations is presented at semantic interpretation of informative model of quality, the functional diagram of realization of text transformations is resulted.*

*Стаття надійшла 20.05.11*