

УДК 655. 28: 681

МЕТАМОВА ОПИСАННЯ МАТЕМАТИЧНИХ ФОРМУЛ

Р.Й. Рінецький

Запропоновано структурний метод описання математичних формул, що ґрунтується на основних засадах математичної лінгвістики. Представлено ієрархічні рівні формул будь-якої складності. Розроблено металінгвістичну мову для описання математичних формул, в якій моделі конструкцій подані з використанням нормальної нотації Бекуса. В результаті більш обґрунтованим стає процес визначення переліку задач та функцій алгоритмів поліграфічного опрацювання математичних формул, створюються передумови для розробки оптимальної логічної структури бази даних.

Предложен структурный метод описания математических формул, базирующийся на основных принципах математической лингвистики. Представлены иерархические уровни формул любой сложности. Разработан металингвистический язык для описания математических формул, в котором модели конструкций поданы с использованием нормальной нотации Бекуса. В результате более обоснованным становится процесс определения перечня задач и функций алгоритмов полиграфической обработки математических формул, создаются предпосылки для разработки оптимальной логической базы данных.

Для описання математичної формули, лінійний запис якої являє собою набір команд і текстових фрагментів, запропоновано структурний метод. Суть методу полягає в тому, що для описання та розпізнавання зображень формули розроблено лінгвістичні конструкції, які складаються з "словника ознак" та "граматики". Метод ґрунтується на математичній лінгвістиці, має прикладний характер та орієнтований на машинне розпізнавання зображень. Дозволяє не тільки розпізнавати та відділяти окремо команди і текстові фрагменти, але й встановлювати призначення команд для їх виконання. При такому підході команди характеризуються набором чисел у вигляді кодів.

Структурний підхід дозволив зобразити ієрархічні рівні формули, що дало можливість описати формулу будь-якої складності за допомогою окремих блоків.

Введемо ряд термінів. Частковий формат (ЧФ) — це частина математичного виразу, що має свій зміст і певну математичну назву. Частковий формат може містити символ чи групу символів, що знаходяться між командами або інші часткові формати. Простий частковий формат (ПЧФ) — частковий формат, що не містить інших ЧФ. Вид часткового формату (V) — послідовність математичних символів. Степінь вкладеності (S) часткового формату — ціле число, що на одиницю більше від кількості часткових форматів, в які одночасно входить даний частковий формат.

Назви часткових форматів утворюються від їх математичного змісту: дріб, степінь, індекс, корінь, інтеграл, сума, добуток, системи рівнянь і нерівностей, матриці і визначники.

Кожний частковий формат задається своєю командою. Наведемо основні команди та їх позначення.

Дріб: [Ч] <#> [З] <#> *
 Степінь: [СТ] <#> *
 Індекс: [ІН] <#> *
 Корінь: [КОР] <#> [ПОК] <#> *
 Квадратний корінь: [КОР] <#> *
 Визначений інтеграл: [ІНТ] <#> [НМ] <#> [ВМ] <#> *
 Невизначений інтеграл: [ІНТ] <#> *
 Кратний інтеграл: [ІНТ-N] <#> [НМ] <#> *(N={1, 2, 3})
 Кратний інтеграл: [ІНТ-kN] <#> [НМ] <#> *(N>3)
 Сума (ряд): [СУМ] <#> [НМ] <#> [ВМ] <#> *
 Сума (ряд): [СУМ] <#> *
 Добуток: [ДОБ] <#> [НМ] <#> [ВМ] <#> *
 Добуток: [ДОБ] <#> *
 Границя: [ГР] <#> [НМ] <#> *
 Вектор: [ВЕК] <#> *
 Система рівнянь (нерівностей): [СР] [n×m] <#> [:] <#> [:] <#> [:] <#> ...
 Матриця: [М] [n×m] <#> [:] <#> [:] <#> [:] <#> ... *
 Визначник: [В] [n×m] <#> [:] <#> [:] <#> [:] <#> ... *

Для визначення суті структури формули та формалізації її подання розглянемо певні класи конструкцій, моделі яких подамо в нормальній нотації Бекуса. При такому способі метасинтаксична мова складається з кінцевого числа фраз, які утворюють металінгвістичні конструкції (формули Бекуса). Для їх побудови використовується два універсальних метасимволи «:=» і «|». Перший з них розуміють як "є за означенням", а другий як "або". Для запису математичних формул можна використовувати об'єкти з конструкцій Бекуса. Одним з таких класів об'єктів є основні символи мови описання, а іншим — імена конструкцій мови описання, які називають металінгвістичними змінними. Ця змінна позначається словами, які пояснюють зміст даної конструкції, і взята в спеціальні дужки: «<», «>».

Сама формула Бекуса утворена за таким правилом. У лівій частині записується металінгвістична змінна, що позначає відповідну конструкцію. Далі йде метасимвол «:=», який з'єднує ліву і праву частини формули. У правій частині вказуються один або кілька варіантів побудови конструкцій. Кожний з варіантів являє собою ланцюжок з металінгвістичних і основних символів. Варіанти правої частини формули розділяють метасимволом «|». Конструкції, що повторюються, заключають у фігурні дужки «{», «}».

Для визначення суті структури математичної формули та формалізації її зображення використаємо введені раніше інформаційні рівні: символ, слово, рядок, простий частковий формат, формульний фрагмент, частковий формат, формула. В кожному з них розглянемо певні класи конструкцій, моделі яких подано в нормальній нотації Бекуса.

<Символ> ::= <буква> | <цифра> | <знак> | <порожній символ>
<Буква> ::= A|B|...|b|a|b|...|b|A|B|...|Z|a|b|...|z

В систему можуть входити також букви інших алфавітів, наприклад грецького.

<Цифра> ::= 0|1|2|3|4|5|6|7|8|9
<Знак> ::= +|-|:|×|÷|Σ|∫|=|<|>|...
<Порожній символ> ::= символ нульової довжини
<Слово> ::= <границя слова> <символ> <границя слова> | <границя слова> {<символ>}
<границя слова>
<Границя слова> ::= <проміжок> | <тире> | <знак>
<Рядок> ::= <слово> | {<слово>}
<ПЧФ> ::= < команда > <рядок> <команда> | < команда > {< рядок > < команда >}
<Команда> ::= <мнемоніка>
<Формульний фрагмент> ::= < рядок > <ПЧФ> <рядок> | < рядок > <ЧФ> <рядок>
<ЧФ> ::= <команда> <формульний фрагмент> <команда> | < команда > {<формульний фрагмент> < команда > }.

Важливою особливістю металінгвістичних формул є наявність в них рекурсії, тобто для опису деяких конструкцій повторно використовуються ці самі конструкції. Наявність рекурсії дещо ускладнює читання та засвоєння металінгвістичних формул, але рекурсії потрібні для того, щоб мова опису формул могла розвиватись.

Аналогічним способом можна розписати інші частини математичної формули. Наведений формалізований опис показує, що кожний з нижчих структурних рівнів є складовою частиною структурного елемента вищого рівня. Глибина опису структур максимальна і доведена до елементарних одиниць. Остаточна формалізація математичної формули матиме такий вигляд:

<Математична формула> ::= [MA]<рядок> {ЧФ <рядок> }@|
[MS]<рядок> {ЧФ <рядок> }@.

Покажемо можливі варіанти часткових форматів математичної формули:

<ЧФ> ::= <ЧФ1> | <ЧФ2> | <ЧФ3> | <ЧФ4> | <ЧФ5> | <ЧФ6> | <ЧФ7> | <ЧФ8> |
<ЧФ9> | <ЧФ10> | <ЧФ11> | <ЧФ12> | <ЧФ13> | <ЧФ14> | <ЧФ15> | <ЧФ16> |
<ЧФ17>.

Їх конструкції залежать від математичного змісту.
Однорядкова математична формула:

$\langle \text{ЧФ1} \rangle ::= \langle \text{рядок} \rangle$
 Дріб:
 $\langle \text{ЧФ2} \rangle ::= [\text{Ч}] \langle \text{ФФ} \rangle [\text{3}] \langle \text{ФФ} \rangle *$
 Степінь:
 $\langle \text{ЧФ3} \rangle ::= [\text{СТ}] \langle \text{ФФ} \rangle *$
 Індекс:
 $\langle \text{ЧФ4} \rangle ::= [\text{ІН}] \langle \text{ФФ} \rangle *$
 Корінь:
 $\langle \text{ЧФ5} \rangle ::= [\text{КОР}] \langle \text{ФФ} \rangle [\text{ПОК}] \langle \text{ФФ} \rangle *$
 Квадратний корінь:
 $\langle \text{ЧФ6} \rangle ::= [\text{КОР}] \langle \text{ФФ} \rangle *$
 Визначений інтеграл:
 $\langle \text{ЧФ7} \rangle ::= [\text{ІНТ}] \langle \text{ФФ} \rangle [\text{НМ}] \langle \text{ФФ} \rangle [\text{ВМ}] \langle \text{ФФ} \rangle *$
 Невизначений інтеграл:
 $\langle \text{ЧФ8} \rangle ::= [\text{ІНТ}] \langle \text{ФФ} \rangle *$
 Кратний інтеграл:
 $\langle \text{ЧФ9} \rangle ::= [\text{ІНТ}-\text{N}] \langle \text{ФФ} \rangle [\text{НМ}] \langle \text{ФФ} \rangle * (\text{N}=\{1,2,3\})$
 Кратний інтеграл:
 $\langle \text{ЧФ9} \rangle ::= [\text{ІНТ}-\text{N}] \langle \text{ФФ} \rangle [\text{НМ}] \langle \text{ФФ} \rangle * (\text{N} > 3)$
 Сума:
 $\langle \text{ЧФ10} \rangle ::= [\text{СУМ}] \langle \text{ФФ} \rangle [\text{НМ}] \langle \text{ФФ} \rangle [\text{ВМ}] \langle \text{ФФ} \rangle *$
 Добуток:
 $\langle \text{ЧФ11} \rangle ::= [\text{ДОБ}] \langle \text{ФФ} \rangle [\text{НМ}] \langle \text{ФФ} \rangle [\text{ВМ}] \langle \text{ФФ} \rangle *$
 Границя:
 $\langle \text{ЧФ12} \rangle ::= [\text{ГР}] \langle \text{ФФ} \rangle [\text{НМ}] \langle \text{ФФ} \rangle *$
 Вектор:
 $\langle \text{ЧФ13} \rangle ::= [\text{ВЕК}] \langle \text{ФФ} \rangle *$
 Система рівнянь (нерівностей): $[\text{СР}] [\text{n} \times \text{m}] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle \dots *$
 Матриця: $[\text{М}] [\text{n} \times \text{m}] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle \dots *$
 Визначник: $[\text{В}] [\text{n} \times \text{m}] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle [;] \langle \text{ФФ} \rangle \dots *$

Металінгвістичні формули дозволяють, по-перше, формально подати основні конструкції мови для написання математичних формул, по-друге, включають в себе елементи семантики, тому що до них входить зміст назви конструкції – послідовність слів у кутових дужках (« $<$ », « $>$ »), що відповідають металінгвістичним змінним.

Запропонований метод формалізації дозволяє оцінити перелік задач поліграфічного опрацювання математичних формул, більш обґрунтовано визначити функції алгоритмів і зв'язки між ними. Полегшує процес проектування машинних інформаційних структур – основних масивів бази даних.