

УДК 004.94

## КОНЦЕПЦІЯ ВИКОРИСТАННЯ ОНТОЛОГІЙ ДЛЯ КОНСТРУЮВАННЯ ЗАСОБІВ ІНДУКТИВНОГО МОДЕЛЮВАННЯ

Г.А. Піднебесна

*Міжнародний науково-навчальний центр інформаційних  
технологій та систем НАН та МОН України, м. Київ*

*pidnebesna@irtc.org.ua*

В роботі розглядається онтологічний підхід до структурування знань при конструюванні засобів індуктивного моделювання.

*Ключові слова: індуктивне моделювання, МГУА, онтологія, декларативне програмування.*

This paper considers the ontological approach to structuring knowledge in the design of inductive modeling tools.

*Keywords: inductive modeling, GMDH, ontology, declarative programming*

В работе рассматривается онтологический подход к структурированию знаний при конструировании средств индуктивного моделирования.

*Ключевые слова: индуктивное моделирование, МГУА, онтология, декларативное программирование*

**Вступ.** Проектування та розробка інтелектуальних комп'ютерних технологій на основі створення високопродуктивних обчислювальних засобів індуктивного моделювання і прогнозування з метою істотного розширення можливостей побудови моделей складних процесів різної природи залишається актуальним напрямком досліджень. Але значна складність проектування та розробки такого програмного забезпечення полягає в тому, що розробник повинен уявляти (знати) множину всіх можливих етапів та варіантів процесів при розв'язанні задачі (екстенціонал задачі), з урахуванням різних допустимих вхідних даних. Ці знання повинні бути втілені у вигляді програмного продукту.

Також значні складності виникають при модифікації програми: програмісту необхідно точно уявляти, які зміни в процесі розв'язування задачі призведуть до очікуваних змін результатів, бути певним в їхній коректності та знати, якими саме змінами в програмі цього можна досягти. Велика вірогідність помилки робить це завдання надзвичайно непростим і вимагає участі у проекті експертів високого класу, що робить виконання таких робіт коштовними. Попередня формальна специфікація дозволяє значно зменшити тривалість та коштовність робіт, оскільки практична користь полягає в полегшенні розуміння вимог, а відповідно, зменшення помилок при реалізації.

Одним з сучасних підходів для спрощення розробки та модифікації програмних продуктів, полегшення їх розуміння, є так звана парадигма

декларативного програмування, яка включає в себе функціональне, логічне та онтологічне програмування.

## **1. Онтологічний підхід як один з різновидів декларативного програмування**

В декларативному програмуванні код має описувати саму задачу, а не один із способів її розв'язання.

*Функціональне програмування* має справу з обчисленнями і використовує функції для виконання обчислень. Цікава особливість функціональних мов програмування полягає в тому, що оскільки вони є реалізацією формальної системи та піддаються математичному аналізу, всі математичні операції, які могли б бути зроблені на папері, також можуть бути застосовані до програм, написаних засобами такої мови. Ці властивості можна використовувати для доведення коректності написаного коду програми, створювати інструменти, які автоматично аналізують код.

*Логічне програмування* засноване на теорії та апараті математичної логіки. Логічні мови програмування зазвичай визначають що саме треба обчислити, властивості самої задачі, а не алгоритм її розв'язання. Програма задається не послідовністю інструкцій, а набором логічних тверджень, множиною фактів та правил, з чого отримуються потрібні результати.

*Онтологічне програмування* – це різновид декларативного програмування, яке базується на роботі з онтологіями. Основна ідея такого підходу полягає в тому, що внутрішня структура програми є відображенням певних знань, які потрібно виразити явно. [1]

Використання онтологій може полегшити описання задачі проектування складних систем з компонентів відповідно з вимогами специфікації та впровадити програму, яка робить таку конфігурацію незалежною від продукту і самих компонентів.

## **2. Поняття онтології в галузі інформаційних технологій**

*Онтологія* – це ієрархічно структурована множина термінів та правил їх використання, що описують предметну галузь. Її можна розглядати як систему множини понять (термінів, концептів) та множини тверджень про ці поняття, на основі яких можна описувати класи, відношення та функції [2]. Онтологія може бути використана за початкову структуру для конструювання відповідної бази знань при розробці інтелектуальних засобів моделювання, в тому числі індуктивного, з використанням МГУА [3].

Часто для формалізації основних понять предметної галузі та відношень між ними розробляються стандартні онтології. Вони використовуються для аналізу

домену, визначають загальні словники (тезауруси) для спільного використання спеціалістами конкретних предметних галузей та для анотування інформації в своїй галузі, полегшують можливість повторного її використання, розширення за рахунок інтегрування декількох вже існуючих онтологій, які стосуються частин великої предметної галузі. [4].

Поняття онтології в області інформаційних технологій було застосовано Томом Грубером у 1993 р. і визначено як «експліцитна специфікація концептуалізації». Тобто, онтологія є явним описом (подібно до формального опису (специфікації) програми) понять та відношень, які можуть існувати для агента або спільноти агентів [5].

Під концептуалізацією, пов'язаною з введенням абстрактних об'єктів для описання навколишнього світу, розуміється абстрактний, спрощений погляд на світ, представлений для досягнення конкретної мети, наприклад, формалізованого представлення ситуацій та підтримки прийняття рішень.

Іншими словами Т.А. Гаврилова в [6] визначає онтологію як точну специфікацію деякої галузі, яка містить словник термінів предметної галузі та множину зв'язків (наприклад, «елмент-клас», «частина-ціле» тощо), які описують співвідношення цих термінів між собою. Це фактично складає ієрархічний понятійний скелет предметної галузі.

Метою створення та використання онтологій можна назвати забезпечення підтримки діяльності з накопичення, розподілення та повторного використання знань в конкретній предметній галузі. Виходячи з цього [2] вводяться критерії, яким повинна задовольняти онтологія:

- Прозорість – онтологія повинна ефективно передавати значення визначеного терміна, необхідного для опису ситуацій.
- Зв'язність – тобто онтологія повинна давати можливість робити висновки, які узгоджуються з початковими визначеннями понять. Принаймні, аксіоми не повинні викликати протиріччя.
- Незалежність від синтаксису – онтологія повинна бути специфікована на рівні знання максимально незалежно від представлення понять на рівні символів.
- Мінімальний базис при високій виразності – онтологія повинна вводити мінімальний базовий набір понять, але достатній для моделювання в необхідних цілях та опису складних ситуацій.
- Зручність для користувача – онтологія повинна дозволяти представляти знання в звичному користувачеві вигляді, бути зрозумілою, осяжною та зв'язаною.
- Ефективність машинної обробки – важлива можливість формалізації онтологій до вигляду, який дозволяє ефективну комп'ютерну обробку, щоб зняти з користувача рутинні операції по підтримці, використанню та розвитку онтологій.

Як правило під формальною моделлю онтології  $O$  розуміють впорядковану трійку вигляду

$$O = \langle C, R, F \rangle,$$

де  $C$  – скінченна (не пуста) множина концептів (понять) предметної області, яку визначає онтологія;

$R$  – скінченна множина відношень між концептами предметної області,

$F$  – скінченна множина функцій інтерпретації (аксіоматизація), заданих на концептах та/або відношеннях онтології  $O$  [7].

Множини  $R$  та  $F$  можуть бути порожніми, що відповідає таким окремим типам онтологій як словник або таксономія понять, тезаурус.

Тобто, онтології повинні перш за все забезпечувати словник понять предметної галузі та множину відношень між цими поняттями.

### 3. Онтологічний підхід в моделюванні предметної галузі індуктивного моделювання

Описуючи формальну модель онтологічної системи [6] виділяють три основних принципи класифікації онтологій:

1. за рівнем формальності: *онтологія верхнього рівня - метаонтологія* – оперує загальними концептами (поняттями) та відношеннями, які не залежать від предметної галузі;
2. за змістом: *предметна онтологія* або *онтологія домену* – складається з понять, що описують конкретну предметну галузь, значущі в ній відношення, множину цих понять та відношень (процедурні та декларативні);
3. за метою побудови: *онтологія задач* або *прикладна онтологія* – поняттями є типи задач, що розв'язуються, а відношеннями є декомпозиція цих задач на підзадачі.

Онтології верхнього рівня можуть бути подібні між собою, оскільки містять знання, загальні для декількох або навіть багатьох предметних галузей, а відповідно, оперують однаковими поняттями як то: сутність, явище, процес, об'єкт, дія, простір, час, подія тощо.

Онтологія предметної галузі використовується як загальний концептуальний базис для ефективного компактного представлення системи знань в конкретній предметній галузі, узагальнює поняття в деяких задачах домену, але відокремлено від самих конкретних задач. Наприклад, онтологія моделювання за статистичними даними може містити концепти: *попередня обробка даних, клас задач, клас моделей, генератор структур, оцінювач параметрів, критерій селекції моделей* тощо. Вони відносяться до всіх методів моделювання незалежно від конкретної задачі.

Онтологія задач специфікує об'єкти діяльності, стану та необхідні для виконання завдань зв'язків серед об'єктів [8]. Частиною такої онтології можуть бути формули для обчислення значень об'єктів.

Онтологія задач необхідна для опрацювання функціональності додатків, які вирішують конкретні типи завдань. Вона використовується при розробці програмного забезпечення для виконання конкретної задачі. Наприклад, в рамках індуктивного моделювання, а саме МГУА [9], розглядається цілий спектр задач, для яких мають бути розроблені прикладні онтології:

- побудова математичної моделі об'єкта (процесу);
- прогнозування перебігу процесу, заданого часовим рядом;
- класифікація – побудова правила віднесення об'єкта до того чи іншого класу за заданими ознаками;
- розпізнавання образів з навчанням (виявлення ефективних ознак та правила розрізнення заданих класів);
- кластеризація (самонавчання, або виявлення ефективних ознак, класів та правил їх розрізнення), в МГУА така задача називається «Об'єктивна Комп'ютерна Кластеризація» (ОКК);
- об'єктивний системний аналіз (ОСА), коли необхідно з'ясувати, які серед вимірюваних змінних є незалежними (вхідними), які залежними (вихідними), які зайвими (неінформативними), і побудувати відповідну модель тощо.

Для формалізації знань в онтології предметної галузі необхідно вибрати базис, в якому будуть описуватись концепти. Прикладом базису може бути набір компонент:

- Класи;
- Відношення – представляють тип зв'язків між концептами ПО, формально можуть бути представлені як підмножина декартового додатку  $n$  множин таких, що  $R: C_1 \times C_2 \times \dots \times C_n$ . (приклади простих бінарних відношень: «бути частиною», «пов'язаний-з» тощо).
- Функції – окремий випадок відношень, коли  $n$ -й елемент визначається по  $(n-1)$  значенням попередніх елементів:  
$$F: C_1 \times C_2 \times \dots \times C_{n-1} \Rightarrow C_n .$$
- Аксиоми – моделюють завжди істинні вирази.

Розглянемо докладніше деякі аспекти структуризації галузі індуктивного моделювання.

Кожен з методів індуктивного моделювання (структурної ідентифікації) може бути охарактеризований такими основними компонентами:

- 1) клас моделей (базисних функцій), для якого він призначений;
- 2) генератор структур у обраному класі базисних функцій;
- 3) метод оцінювання параметрів генерованих структур моделей;
- 4) критерій селекції або якості моделей.

Розглядаючи кожен з методів з точки зору таких складових, можна утворити множини класів моделей, генераторів структур, методів оцінювання параметрів і критеріїв вибору моделей і аналізувати ці множини як самостійні наукові об'єкти.

В результаті структуризації знань з конкретної проблеми визначаються принципи формування алгоритмічних модулів, узагальнюючих функції різних методів розв'язання конкретного класу задач в єдиний багатофункціональний алгоритм. Він характеризується певним набором ключових параметрів. Різні комбінації цих параметрів утворюють окремі методи як окремі випадки.

Нехай

$KM$  – множина класів моделей,  $KM = \{k_i\}, i = \overline{1, K}$ ;

$GS$  – множина генераторів структур моделей,  $GS = \{g_j\}, j = \overline{1, G}$ ;

$OP$  – множина методів оцінювання параметрів структур,  
 $OP = \{p_q\}, q = \overline{1, Q}$ ;

$CR$  – множина критеріїв якості моделей,  $CR = \{r_t\}, t = \overline{1, R}$ .

Тоді множину методів побудови моделей  $L$  можна представити у вигляді прямого добутку множин:

$$L = KM \times GS \times OP \times CR.$$

Під алгоритмом будемо розуміти певний елемент множини  $L$ :

$$l_d = \{k_i, g_j, p_q, r_t\},$$

$$i = \overline{1, K}, j = \overline{1, G}, q = \overline{1, Q}, t = \overline{1, R}, d = \overline{1, K \times G \times Q \times R}.$$

Набір  $\langle KM, GS, OP, CR \rangle$  можна розглядати як початковий базис для побудови онтологій індуктивного моделювання.

#### 4. Приклади етапів побудови онтологій галузі індуктивного моделювання

Найчастіше онтологія будується у вигляді дерева або семантичної мережі, яка складається з концептів та зв'язків між ними.

При побудові онтології предметної галузі одним з найважливіших кроків вважається визначення множини концептів. Перш за все будується словник термінів, властивих предметній галузі. Наприклад, фрагментом словника з індуктивного моделювання може бути набір понять: клас моделей, генератор структур, оцінювання параметрів, критерій незміщеності, критерій регулярності, лінійна регресія, авторегресія, комбінаторний алгоритм, багаторядний алгоритм, апроксимація, виділення тренду, прогнозування, комбінаторно-селекційний алгоритм, розбиття вибірки, перевірна вибірка, екзаменаційна вибірка тощо.

Далі виконується ранжування списку понять за принципом «вище-нижче»:  
множина генераторів структур моделей;  
перебірні генератори; ітераційні генератори;  
генератори повного перебору; генератори спрямованого перебору; багаторядні ітераційні; релаксаційні;  
тощо...

Наступним кроком будується множина відношень, яка може складатись з таких елементів як: *ціле\_частина*, *бути\_характеристикою*, *рід\_вид*, *is\_a* тощо.

Згідно з цим можна побудувати ієрархічне дерево, наприклад, для множини класів моделей (рис. 1):

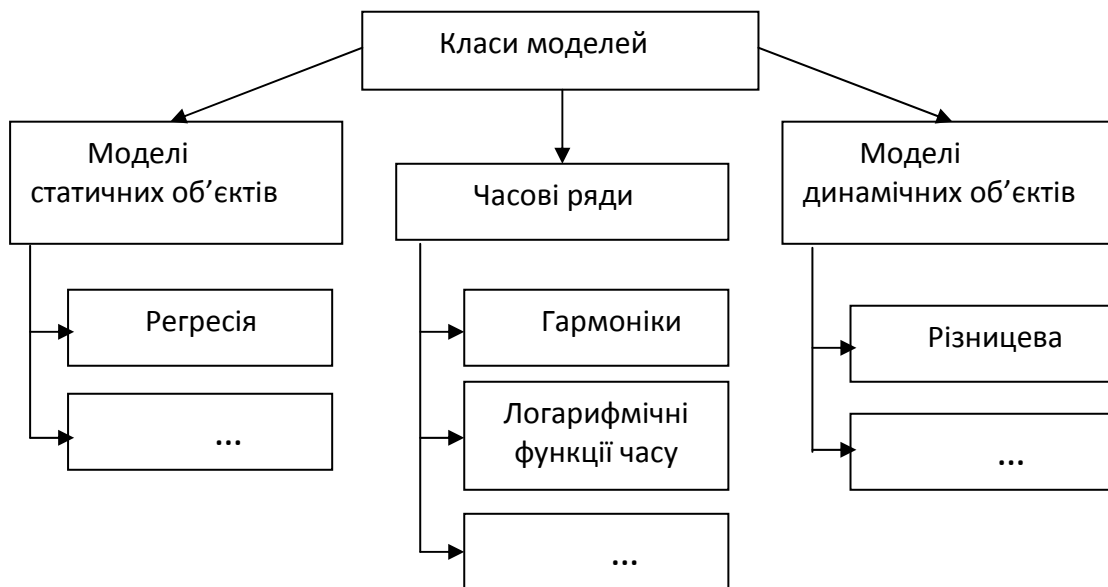


Рис. 1 Фрагмент дерева відношення *елемент\_класу*

## Висновки

При проектуванні інтелектуальних комп'ютерних систем моделювання і прогнозування складних процесів різної природи за статистичними даними актуальним напрямком досліджень є використання онтологічного підходу.

Побудова онтологій предметної галузі індуктивного моделювання дасть можливість істотного спрощення процесу розробки специфікацій та програмних продуктів для розв'язання конкретних задач побудови моделей та прогнозів, а також для розширення можливостей модифікації існуючих комп'ютерних засобів.

Представлено приклади декількох кроків при побудові фрагментів онтологій індуктивного моделювання.

## Література

1. Грибова В.В, Клещев А.С. Онтологическая парадигма программирования. // конференція OSTIS-2012 (Open Semantic Technologies for Intelligent Systems), 2012. – с. 213-220. (<http://conf.ostis.net>)
2. Скобелев П.О. Онтологии деятельности для ситуационного управления предприятиями в реальном времени. // Ж-л «Онтология проектирования». Самара : Новая техника. - 2012. - №1. - с. 6 – 39.
3. Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования. - Киев: Наук. думка, 1985. – 216 с.
4. Валькман Ю.Р. Онтологии: формальное и неформальное. // доклад на семинаре «Образный компьютер», 08.11.2011, Киев: МНУЦ ИТС. (<http://image.irtc.org.ua/seminars/archive/ok/ru>).
5. Gruber T. R.. A translation approach to portable ontologies. // Knowledge Acquisition. – 1993. - 5(2). – p.199-220.
6. Гаврилова Т.А. Онтологический подход к управлению знаниями при разработке корпоративных систем автоматизации. // [http://bigc.ru/publications/bigspb/km/ontol\\_podhod\\_to\\_uz.php](http://bigc.ru/publications/bigspb/km/ontol_podhod_to_uz.php)
7. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний в интеллектуальных системах // СПб. - 2000. – 384с.
8. Шалфеева Е.А.. Возможности использования онтологий при разработке и сопровождении программных систем. // Владивосток: ИАПУ ДВО РАН. – 2011.- с. 23.
9. Степашко В.С., Савченко Є.А., Піднебесна Г.А. Індуктивне моделювання як процес послідовного прийняття рішень // Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту» (ISDMCI'2010), Євпаторія. – Херсон: Видавництво ХНТУ, 2010. – Т.2. – С. 457-460.
10. Палагин А.В., Петренко Н.Г., Малахов К.С.. Методика проектирования онтологии предметной области. // Комп'ютерні засоби, мережі та системи. – 2011 - № 10. – с. 5-12. ([http://archive.nbu.gov.ua/portal/natural/Kzms/2011/2011\\_st1.pdf](http://archive.nbu.gov.ua/portal/natural/Kzms/2011/2011_st1.pdf))