

О.В. ПАЛАГІН, М.Г. ПЕТРЕНКО

## **МОДЕЛЬ КАТЕГОРІАЛЬНОГО РІВНЯ МОВНО-ОНТОЛОГІЧНОЇ КАРТИНИ СВІТУ**

**Abstract:** *The new approach to the decision of one of the main tasks of systems of an artificial intellect – the analysis and understanding natural-language texts is offered. The abstract model of a ontological picture of the world which is based on a lexicographic database of a natural language and semantic-syntactical relations between language units is developed. The variant of ontology categories of a conceptual level is described.*

**Key words:** *ontology, a ontological-language picture of the world, semantic-syntactical relations.*

**Анотація:** *Запропоновано новий підхід до розв'язання однієї з головних задач систем штучного інтелекту – аналізу та розуміння природномовних текстів. Розроблено абстрактну модель мовно-онтологічної картини світу, що базується на лексикографічній базі даних природної мови та семантико-синтаксичних відношеннях між мовними одиницями. Описано варіант онтології категорій концептуального рівня.*

**Ключові слова:** *онтологія, мовно-онтологічна картина світу, семантико-синтаксичні відношення.*

**Аннотация:** *Предложен новый подход к решению одной из главных задач систем искусственного интеллекта – анализа и понимания естественно-языковых текстов. Разработана абстрактная модель языково-онтологической картины мира, которая базируется на лексикографической базе данных естественного языка и семантико-синтаксических отношениях между языковыми единицами. Описан вариант онтологии категорий концептуального уровня.*

**Ключевые слова:** *онтология, языково-онтологическая картина мира, семантико-синтаксические отношения.*

### **1. Вступ**

На сучасному етапі розвитку computer science при розробці інтелектуальних інформаційних систем (ІС) загальноновизнано включати до їх складу процедури аналізу та розуміння природномовних текстів (ПМТ) на основі онтологій. Про актуальність такого включення свідчить зростання наукових праць з проблематики розробки концептуальних моделей побудови та опрацювання онтологічних структур і баз знань з відповідних проблемних областей [1–8, 14–17, 19]. Практична значимість вказаних розробок, в основному, залежить від повноти інтерпретаційних моделей семантичних структур ПМТ та їх формального представлення в комп'ютерно-орієнтованій ІС. Під повнотою ми розуміємо включення в модель як складової семантики першого ступеня (або об'єктової складової), так і складової другого ступеня (або акторної складової, від англійського слова *actor* – діючий об'єкт, суб'єкт). Такий розподіл семантики добре узгоджується як з онтологічною ієрархією концептуальних категорій, так і з складністю виконання обчислювальних процедур при комп'ютерній обробці ПМТ.

З точки зору лінгвістики, семантична складова першого ступеня описується на рівні граматики окремих частин мови, в той час як складова другого ступеня вже описується синтаксичними структурами таких одиниць синтаксису, як речення, абзац, параграф, розділ і текст. З точки зору математичної логіки, якщо перший ступінь можна описати (досить умовно) численням висловлень, то другий ступінь описується численням предикатів з квантифікованими змінними.

Найбільшої повноти (і відповідно найбільшого ступеня складності) набувають моделі, що описують природномовний текст в цілому. Такі моделі описують, зокрема, деякий сценарій (як найвищу категорію, описуючу форми всього, що знаходиться в постійному русі), який відтворює зміст ПМТ. У свою чергу, як ПМТ підрозділяється на синтаксичні одиниці, так і загальний сценарій підрозділяється на окремі сценарії, ситуації та елементарні ситуації.

Розробка концептуальної моделі адекватного представлення морфології, синтаксису та семантики обох рівнів ПМТ є досить складною науковою задачею. На шляху до її розв'язання необхідно:

- розробити єдину методологію побудови як загальної (мета)онтології, так і онтологій з конкретних предметних областей;
- розробити онтологічне представлення мовної картини світу (МКС) з орієнтацією на його інтерпретацію засобами сучасних комп'ютерних технологій;
- розробити метамову уніфікованого представлення як онтологічної структури ПМТ, що опрацьовується, так і онтологічної структури МКС;
- розв'язати питання ефективного представлення та використання у процесі інтерпретації ПМТ т.з. “фонових” знань, аксіом та фактів, імпліцитно присутніх і безпосередньо впливаючих на онтологічну структуру природномовного повідомлення;
- формалізувати процедуру узагальнення, тобто переходу від нижніх рівнів до вищих в ієрархії онтологічних категорій;
- розширити можливості проблемно-орієнтованих обчислювальних засобів, які, з одного боку, повинні забезпечити ефективну інтерпретацію онтологічних структур різних рівнів, а з іншого, зменшити складність розробки як системного, так і прикладного програмного забезпечення. Одним із таких засобів може бути програмно-апаратна система на базі сучасних програмованих логічних інтегральних схем (ПЛІС) – технологій.

## 2. Постановка задачі

Однією із основних цілей онтології є представлення структури розходжень, які можуть бути використані для розрізнення та класифікації речей матеріального та нематеріального світу, що існують та визначаються словами, які їх описують. Згідно з цим, наше дослідження не ставить собі за мету розробку концепції побудови онтології взагалі, а його об'єктом є загальнонаукова мовна картина світу та її онтологічне представлення, що разом складають мовно-онтологічну картину світу (МОКС).

Як визначено в [7], основною задачею МКС є фіксація змісту вхідної інформації, причому інформаційні моделі такої фіксації повинні адекватно відтворювати обчислювальні процедури у процесі інтерпретації ПМТ.

Для вирішення цієї задачі комп'ютерно-орієнтована МКС повинна відповідати таким вимогам:

- системної повноти як у значенні множини лексем даної мови, так і у значенні формул їх тлумачення;
- мати у своєму арсеналі інструментарій для семантико-синтаксичного аналізу вхідного тексту і, передусім, рішення проблеми потенційної багатозначності вираження змісту вхідного тексту;
- надавати можливості використання як семіотичної основи для фіксації змісту вхідної інформації;
- виконувати функції мовно-категоріальної надбудови баз знань у конкретних предметних областях.

На основі переліку основних функцій в [7] наведено таке визначення МКС.

МКС – це замкнута, формалізована інформаційна (лексикографічна) система, що розвивається, і яка забезпечує семантичне тлумачення елементарних одиниць природної мови та їх устояних конструкцій різної складності у процесі аналізу та інтеграції вхідної інформації, представленої природною мовою.

Виходячи з наведених вище функцій МКС та її онтологічного представлення, МОКС можна визначити як один із центральних компонентів знання-орієнтованої онтолого-керованої інформаційної системи, інструментарій для створення відповідних систем в проблемних областях. В неї включено тільки загальнонаукові, загальнонавчальні (або енциклопедичні) поняття, тобто основні лексичні одиниці Тлумачного словника української мови. При цьому множини мовних одиниць частин мови структуруються окремо, відповідно зі своїми сукупностями базових відношень (для повнозначних частин мови).

Однією з найважливіших функцій МОКС вважається ефективно формалізоване представлення синтаксичних одиниць природної мови. В [9] окреслено, що найважливішим виглядає завдання розробки ефективних методів формалізації, які можуть бути застосовані при розв'язанні основної проблеми мовознавства – проблеми значення одиниць та конструкцій мови. Причому, формалізми повинні бути такими, щоб результати представлялися в уніфікованій, операціонально орієнтованій формі, доступній і зручній для застосування в комп'ютерних технологіях.

В загальноприйнятому варіанті процедури аналізу та розуміння ПМТ виконуються лінгвістичним процесором (ЛП) і, як правило, на програмному рівні і тільки для об'єктової складової семантики (хоча інколи дещо розширеної) [12]. Описана вище суттєва різниця між об'єктовою та акторною семантиками, особливо в їх складності інтерпретаційних моделей, спонукала нас ввести для інтерпретації останньої семантичний процесор (СП), окрім згаданого вже лінгвістичного.

Загальний алгоритм, що реалізується за допомогою ЛП, включає такі традиційні етапи аналізу, як морфологічний, синтаксичний та об'єктово-семантичний. А особливістю лінгвістичної обробки в такій підсистемі є підпорядкування цих етапів заключному процесу формування елементів формалізованого представлення знань за допомогою засобів деякої метамови, що однаково добре описує онтологічні структури різних рівнів.

Основним призначенням СП є побудова формалізованого опису вхідного ПМТ та його відображення в онтологічному дереві мовної картини світу. Іншими словами, головною задачею СП [7] є відображення структури тексту на онтологічну структуру МОКС і фіксація семантико-синтаксичної структури окремих речень і текстових фрагментів у вигляді відповідних сукупностей індексів, що зв'язують відношеннями повну множину лексем та їх значень, представлених в МОКС.

### **3. Побудова онтології МОКС**

У багатьох працях з питань розбудови інтелектуальних інформаційних систем підкреслюється важливе значення побудови ієрархічних структур категорій верхнього рівня. Так, в [4] відмічено, що побудова системи категорій є необхідний етап розробки інтелектуальних систем, заснованих на знаннях, а також систем, що забезпечують автоматизацію розуміння смислу тексту. У [8] висловлено дещо іншу думку: "Суттєвим є побудова мовно-онтологічної картини світу (МОКС) як

основи інформаційної взаємодії систем різної матеріальної природи". Там же наведено рисунок піраміди-надбудови МОКС над прямокутниками – предметними областями. Ґрунтова наукова розробка мовної картини світу та ієрархії сем викладена у праці [10], де модель семантичних відношень у лексиці набуває схеми смислового каркасу, на якому базується глобальна "сферична" мережа різнотипних семантичних відношень. Смиловий каркас побудовано за двома параметрами: гносеологічним – категорії пізнання (буття, простір, час, рух, окреме, якість, кількість, відношення) та онтологічним – сфери буття (людина, природа, суспільство). Аналогічний приклад розбудови лексико-семантичних відношень першого і другого ступеней наведено у праці [11].

Схеми структуризації МОКС, наведені в цих працях та ряді інших, по-перше, страждають певною суб'єктивністю підходу до систематизації, структуризації та класифікації понять; по-друге, неповнотою лексичного континууму та відношень між представленими лексичними одиницями (розуміється, передусім, не їх теоретичний доробок, а комп'ютерне опрацювання, точніше – програмне моделювання); по-третє, вони відокремлені від сучасних інтелектуальних інформаційних технологій та їх застосування.

Стосовно нашого дослідження:

- за основу взято ієрархічну структуру категорій вищого рівня, синтезовану на логіко-філософських засадах [19];

- повнота лексичного континууму забезпечується відтворенням у відповідних таблицях лексикографічної бази даних повної множини лексем Тлумачного словника української мови;

- дослідження як таке орієнтовано на розробку апаратних засобів підтримки (семантичного процесора) на базі сучасних технологій.

Розглянемо співвіднесеність понять мовної картини світу, онтології та загальнонаукових баз знань. Схематично їх взаємодію представлено на рис. 1.

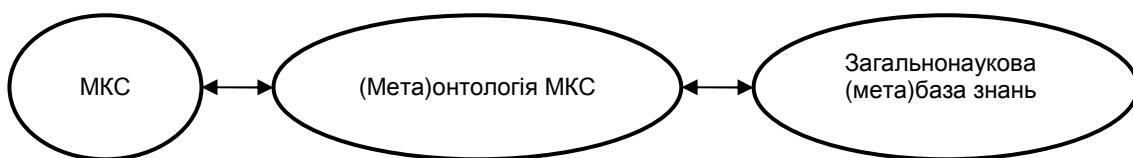


Рис. 1. Співвіднесеність загальнонаукових та онтологічних понять

МОКС є основним вихідним та підготовчим матеріалом для комп'ютерного опрацювання ПМТ. Вона вміщує інформаційно-лінгвістичну та програмно-апаратну складові.

Відповідно, інформаційно-лінгвістична складова включає множину словникових таблиць, закріплених за всіма частинами мови (включаючи і підтаблиці); таблиць формування словозмінної для лексем змінних частин мови; бази знань синтаксичних зв'язків сполучення слів різних частин мови у реченні.

Програмно-апаратна складова представляє собою ЛП з функцією інтелектуального аналізу вхідного ПМТ, організованою за реляційною моделлю вихідної інформації та опрацювання вхідного ПМТ на рівні об'єктової семантики.

Одним із головних в мовно-онтологічній інформаційній системі (МОІС) є блок, який відтворює (мета)онтологію МОКС. Його основною функцією є побудова формалізованого опису

відображення вхідного ПМТ на ієрархічну структуру МОКС з відтворенням повної акторної складової семантики, присутньої в ньому.

Таке відображення можна записати у вигляді

$$G : C_o(\text{ПМТ}) \rightarrow C_A(\text{ПМТ}), \quad (1)$$

де  $C_o(\text{ПМТ}) = \bigcup_{i=1}^N C_{\text{Л}i}$ , де  $\tilde{N}_{\text{Е}i}$  – формалізований опис  $i$ -ої компоненти об'єктової

складової семантики слів повнозначних частин мови, що входять у вихідний ПМТ;  $i = \overline{1, N}$  – кількість вказаних слів;  $C_A(\text{ПМТ}) = F_A(C_o)$  – формалізований опис акторної складової семантики вихідного ПМТ.

Програмно-апаратну частину МОІС названо семантичним процесором. Задачі користувача в повному обсязі вирішуються при взаємодії ЛП, СП та процесора інтерпретації бази (мета)знань і бази знань з конкретної предметної області (побудова останнього представляє окрему задачу).

Уявлення про взаємозв'язок між загальнонауковою базою знань і онтологією МОКС, висловлюючись мовою метафори, можна подати як "квітуче дерево з повним складом листя і гілок в літній період" для бази знань, і "тим же самим деревом ранньою весною і після його обробки садівником" – для онтології.

Отже, ми підійшли до визначення онтології МОКС та формулювання відповідно до нього переліку задач, які потрібно вирішити при розбудові такої онтології.

*Онтологія МОКС – це відкрита, експліцитно задана на лексико-змістовому континуумі лексикографічна система, в якій сукупність категоріальних понять високого рівня формально обґрунтована та впорядкована у складну ієрархічну структуру за основними типами лексико-семантичних відношень.*

Онтологія МОКС (О) як формальний опис загальнонаукових, енциклопедичних знань представляється у вигляді [8]

$$O = \langle X, R, F \rangle, \quad (2)$$

де  $X$  – кінцева множина концептів, виражених лексичними одиницями верхнього рівня природної мови, причому  $X \subseteq U$  і непуста;  $U$  – повна множина лексичних одиниць природної мови;  $R$  – кінцева множина лексико-семантичних відношень між концептами, основними з яких є відношення (в першу чергу, за такими видами абстракції, як класифікація, узагальнення, агрегація та асоціація) "рід-вид", "частина-ціле", "елемент-клас" та "об'єкт-атрибут";  $F$  – множина функцій інтерпретації, заданих на концептах і/чи відношеннях.

Загальні принципи та обґрунтування розбудови ієрархічної структури будь-якої онтології описані в багатьох працях, наприклад, [4–8, 10, 19]. Основний зміст цих описів можна звести до таких положень [4].

1. Система понять відображає структуру реальної дійсності (будь-якої предметної області) засобами ієрархії концептуальних та категоріальних відношень.

2. Ієрархічна структура реальної дійсності припускає, що існує вершина цієї структури. Причому, із усіх теоретично можливих варіантів структури дійсності на вершині ієрархії реальних об'єктів знаходиться один гранично широкий клас об'єктів.

3. Вказана властивість структури системи понять абстрактного мислення повинна бути доповнена родовими властивостями, або атрибутами.

4. Система понять організується в ієрархічну структуру згідно з концептуальними відношеннями для відповідних рівнів. Існують різні, близькі за змістом варіанти класифікацій семантичних відношень [11, 18]. Наведені нижче класи відношень відповідним чином структуровані для вищих рівнів ієрархії.

- Відношення класифікації.
- Ознакові відношення.
- Кількісні відношення.
- Якісні відношення.
- Відношення порівняння.
- Відношення приналежності.
- Часові відношення.
- Просторові відношення.
- Каузальні відношення.
- Інструментальні відношення.
- Інформаційні відношення.
- Порядкові відношення.
- Модальні відношення.
- Модифікатори.
- Квантифікатори.

Уточнимо зміст поняття ієрархічного структурування в онтології МОКС відносно лексичного континууму природної мови. Смісл будь-якого висловлювання виникає, коли текстові компоненти вступають у певні дійові відношення (об'єкт → дія). Сфера дії визначається кількістю об'єктів, які вона охоплює.

Такий розгляд дуже спрощено описує структуру відображення формули (1) і більше орієнтований на опис об'єктової складової семантики. Очевидно, що для повного відображення побудованих лінгвістичним процесором синтаксичних структур вхідного ПМТ, опису акторної складової тексту необхідно враховувати перш за все базові повнозначні частини мови. В узагальненій схемі взаємодії структурних компонентів МОКС за основу приймемо припущення щодо представлення в ній та семантичної інтерпретації відповідних частин мови:

- об'єкти – як іменники;
- дії – як дієслова;
- атрибути об'єктів – як прикметники;
- атрибути дій – як прислівники.

Така інтерпретація добре узгоджується з відомою концепцією Word Net [9], семантично орієнтованою базою даних лексичних одиниць. У ній лексика природної мови розбита на п'ять груп: іменники, дієслова, прикметники, прислівники та слова-функції.

Перші чотири групи та прийменники з п'ятої представляють собою основні складові для опису обох семантичних складових природномовних текстів. Отже і ієрархічна структура МОКС повинна представляти собою деякий гіперграф як композиція чотирьох ациклічних орієнтованих графів ієрархічних структур іменника, дієслова, прикметника та прислівника. Кожен з цих графів будується у відповідності зі своїми концептуальними відношеннями, притаманними кожній із зазначених частин мови. Відображення семантичних особливостей поєднання іменника з прийменником враховується у графі для іменника.

Згідно з наведеним вище описом на рис. 2 представлено абстрактну модель мовно-онтологічної картини світу, де прийнято такі скорочення:

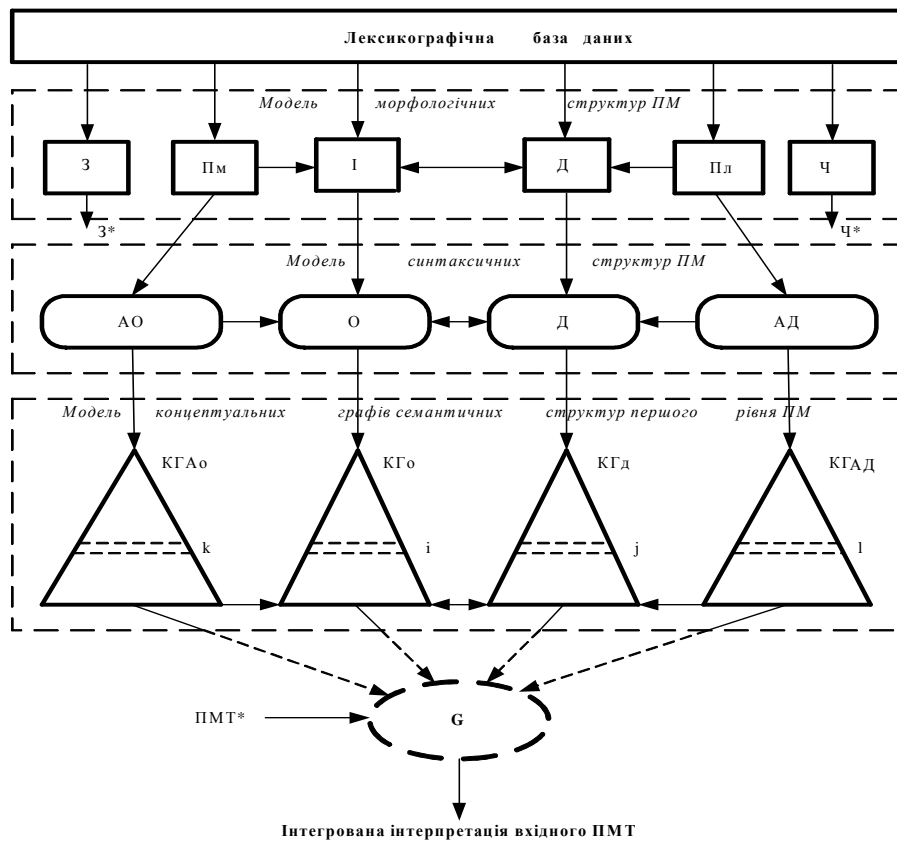


Рис. 2. Абстрактна модель мовно-онтологічної картини світу

- З, Пм, І, Д, Іл, Ч – повнозначні частини мови, відповідно знаменник, прикметник, іменник, дієслово, прислівник та чисельник;
- АО, О, Д, АД – елементи синтаксичних структур, відповідно атрибут об'єкта, об'єкт, дія та атрибут дії;
- КГ<sub>АО</sub>, КГ<sub>О</sub>, КГ<sub>Д</sub> та КГ<sub>АД</sub> – елементи семантичних структур першого ступеня, відповідно концептуальні графи атрибуту об'єкта, об'єкта, дії та атрибуту дії;
- З\*, Ч\* – певним чином впорядковані морфологічні таблиці для частин мови, відповідно знаменник та чисельник, у вихідному ПМТ;

– ПМТ\* – вихідний природномовний текст після інтерпретації модулем обробки семантики першого ступеня.

Відносно представлення абстрактної моделі ще раз зазначимо, що модель морфологічних структур природної мови (ПМ) є спрощеною і не розкриває всіх повнозначних частин мови та взаємовідношень між ними. Таке спрощення є справедливим, тому що в цій моделі зроблено акцент саме на онтологічну складову.

Як було визначено, онтологічний граф не включає поняття нижнього рівня. Але формалізованого критерію визначення межі нижнього рівня не існує і в кожному випадку повністю залежить від інтуїції та професійних поглядів дослідника.

#### 4. Синтез категорій верхнього рівня

Ієрархічна структура на рис. 3 являється результатом концептуалізації понять згідно з усталеними філософськими поглядами (від Геракліта до Пірса і Уайтхеда) [19]. Символ **У** позначає нейтральне, інтуїтивно зрозуміле поняття універсального типу, або Універсуму.

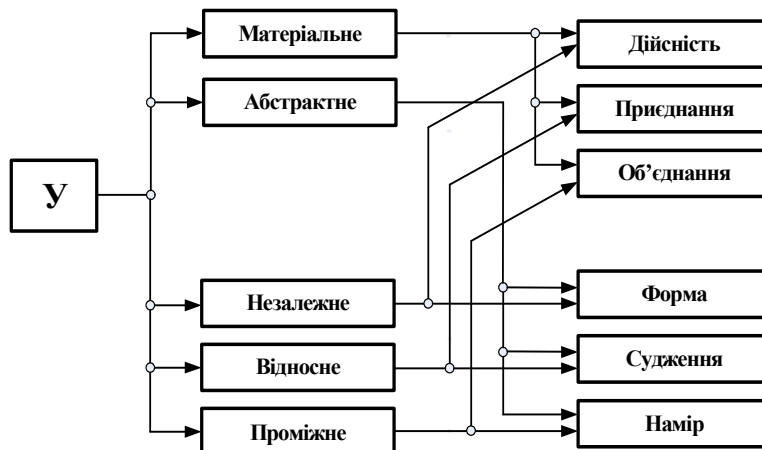


Рис. 3. Ієрархія категорій *Матеріальне* і *Абстрактне* та *Незалежне*, *Відносне* і *Проміжне*

Формально, **У** є типом найвищого (нульового) рівня категоризації, не має ніякої диференціації і якому задовольняють такі аксіоми:

- існує деякий  $x$  такий, що  $(\exists x)U(x)$ ;
- всякий  $x$  є елементом **У**:  $(\forall x)U(x)$ ;
- кожна категорія є підкатегорією **У**:  $(\forall k: \text{Категорія}) k \leq U$ ;
- всі підкатегорії визначаються додатковими відмінностями в **У**, щоб показати, як вони відрізняються від **У** і кожна від іншої.

Нижче **У** розгортається у двонаправлене розбиття між групами категорій: для першої групи – Матеріальне для чого-небудь, складеного з матерії чи енергії, і категорія Абстрактне – для суто інформаційних структур; та для другої групи – Незалежне, Відносне і Проміжне. Третій рівень ділить зазначені групи на тріади згідно з розрізненням Пірса про Первинність, Вторинність і Третинність. Групи категорій: Дійсність, Приєднання та Об'єднання – для Матеріального, як і Форма, Судження та Намір – для Абстрактного є відповідно Первинністю, Вторинністю і Третинністю.



Отже, шість категорій третього рівня сгенеровано перетином двохстороннього та трьохстороннього розрізень, що відповідає принципу комбінування примітивів Лейбніца, який використовується при генерації категорій.

Враховуючи вищевикладене до верхнього рівня **У**, дихотомія Геракліта генерує категорії Матеріальне і Абстрактне, а трихотомія Пірса – три категорії: Незалежне, Відносне і Проміжне.

*Незалежне* – категорія дійсних сутностей і їх абстрактних форм, які характеризуються притаманними якостями або Первинністю. В логіці незалежні сутності можуть бути представлені ім'ям сутності чи одномісним предикатом  $P(x)$ , таким, що описує деякий аспект сутності  $x$ , не беручи до уваги що-небудь зовнішнє до  $x$ .

*Відносне* включає матеріальне Приєднання і абстрактне Судження, які можуть бути представлені двохи́сними предикатами  $P(x, y)$ , і відповідають Вторинності Пірса. Приєднання – це категорія матеріального, яка зв'язує сутність  $x$ , що приєднується, з сутністю  $y$ , що приєднує. Судження – це категорія абстрактного, яка зв'язує форму  $x$  з сутністю  $y$ , описаною за допомогою  $x$ . Наприклад, судження про те, що сонце являється круглим, значить, що форма круга може бути використана для характеристики певної сутності в небі.

*Проміжне* відповідає Третинності Пірса, яке включає матеріальне - Відносне і абстрактне - Намір. Як приклад, можна навести креслення будівлі, який є незалежною сутністю, що включає олівцеві помітки на папері. Креслення як проміжна сутність є планом для підрядчика, який переводить конфігурацію поміток у структуру дерева, металу чи бетону. План  $x$  реалізує підрядчик  $y$  в конструкції будівництва  $z$ . (Це трьохмісне відношення не може бути вираженим кон'юнкцією двохи́сних відношень).

Наведений опис і частково рис. 3 демонструють використання принципу множинного наслідування, за якого кожна категорія нижнього рівня наслідує властивості від двох інших категорій з попереднього рівня.

Виходячи з принципу трихотомії, філософи-логіки розрізняли (і включали в онтологію) третю групу категорій другого рівня (на рис. 3 перша група категорій включає Матеріальне і Абстрактне, друга група – Незалежне, Відносне і Проміжне), а саме – *Тривале* і *Подійне*. Вони були єдині в поглядах, що "всі речі існують у русі, але деякі речі піддаються швидкій зміні, в той час як інші залишаються порівняно сталими" [19]. Щоб розмістити об'єкти в онтології, оснований на процесах, вони відрізняли сталі об'єкти, які мають сталу тотожність більше деякого проміжку часу, від постійно змінних подій, послідовні стадії яких можуть бути не тотожними. Сталий об'єкт називають *Тривалим*, а процес чи подія, які не мають сталих характеристик, називають *Подійним*. Різниця між ними визначається способом, яким сутність ідентифікована:

- *тривалість* має сталі атрибути чи характеристики, які дають можливість у різний час розпізнати її в різних аспектах як ту ж саму індивідуальність;
- *подійність* існує у стані постійної зміни, яка перешкоджає їй бути розпізнаною сталою множиною атрибутів. Замість цього вона може бути ідентифікована тільки своїм розташуванням у деякій системі координат "простір-час".

Класифікація таких сутностей, як *Тривале* чи *Подійне*, залежить від масштабу часу і рівня деталізації з точки зору деякого дослідника.

Завдяки дихотомічній різниці між *Тривалим* та *Подійним*, шість категорій на рис. 3 перетворюються на дванадцять категорій, представлених на рис. 4. Всі вони одержані комбінаціями трьох базових груп відмінностей (чи вимірів) при категоризації Універсального:

*Незалежне - Зв'язне - Проміжне;*

*Матеріальне - Абстрактне;*

*Тривале - Подійне.*

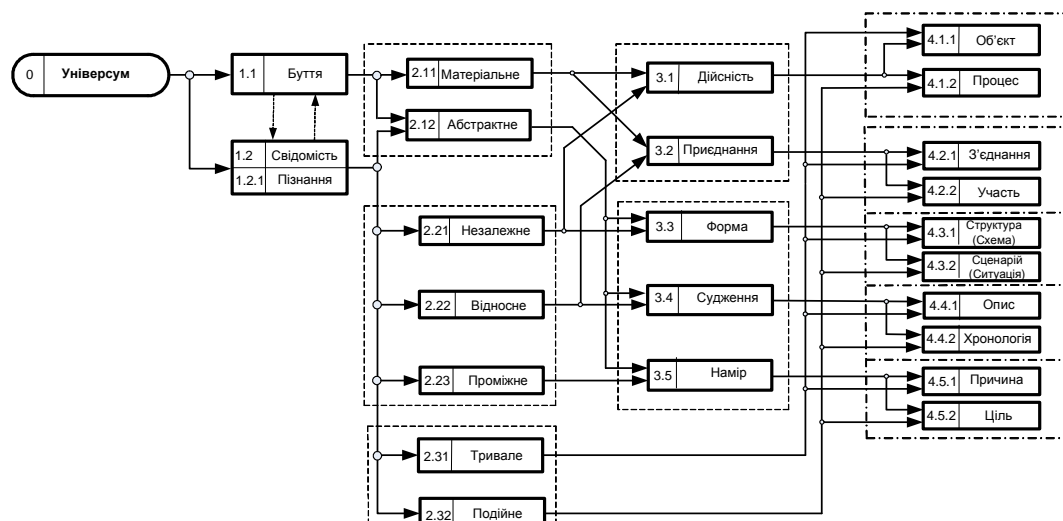


Рис. 4. Ієрархія категорій верхнього рівня

Кожна з інших категорій – синонім для комбінації категорій, з яких вона була одержана: *Об'єкт*, наприклад, може бути представлений акронімом НМТ: Незалежне Матеріальне Тривале.

Слід зазначити, що до вказаних 10-ти категорій на рис. 4 не включені інші категорії, що могли бути сгенеровані за комбінаційним методом Лейбніца. Як указано в [19], включені 10 категорій є найбільш корисними для загальних використань (що відповідає меті нашого дослідження) і являються центральними.

Розглянемо більш детально побудову онтологічної структури для понять – категорій верхнього рівня, які в природній мові репрезентовані іменниками (чи прикметниками, які замінюють іменники). Подану на рис. 4 ієрархічну структуру синтезовано, головним чином, на базі праць [10,11,19].

Самий верхній (нульовий) рівень (порядковий номер рівня вказано зліва у прямокутнику) представлено однією вершиною, яку названо "Універсум" (У).

На першому рівні дві вершини – "Буття" і "Свідомість/Пізнання". Цим підкреслюється різна природа матеріального і абстрактного. На подальших рівнях враховується взаємний перетин дуг, що виходять з вершин другого рівня категорій. Урахування таких перетинів суттєво впливає на практичну значимість отриманих результатів при вирішенні конкретних задач користувачів.

Опишемо включені в ієрархію категорії, відповідно з їх структурними зв'язками, а також наведемо аксіоми (де можливо) для відповідних категорій. Підкреслимо, що кожна категорія в ієрархії наслідуює всі властивості і аксіоми кожної вищевказаної категорії.

Другий рівень представлено сьома вершинами.

2.11. **Матеріальне** (М) – об'єкт, який має місцезнаходження в системі координат простір-час. Для нього можна навести такі аксіоми:

- що-небудь матеріальне  $x$  існує в деякому місці:  $(\forall x:M)(\exists y:\text{Місце})loc(x,y)$ , де  $loc(x,y)$  – відповідний предикат місцезнаходження  $x$  в  $y$ ;

- що-небудь матеріальне  $x$  відбувається в деякий момент часу:  $(\forall x:M)(\exists t:\text{Час})ptime(x,t)$ , де  $ptime(x,t)$  – відповідний предикат відбування  $x$  в  $t$ .

2.12. **Абстрактне** (А) – сутність, для якої виконуються такі аксіоми:

- немає ніякого абстрактного  $x$ , розміщеного у просторі:  $\neg(\exists x:A)(\exists y:\text{Місце})loc(x,y)$ ;
- немає ніякого абстрактного  $x$ , що відбувається в момент часу:  $\neg(\exists x:A)(\exists t:\text{Час})ptime(x,t)$ .

2.21. **Незалежне** (Н) – сутність, яка характеризується деякою природною первинністю, незалежною від будь-яких взаємовідношень. Для неї відношення " $x$  має  $y$ " не виконується, тобто  $(\forall x:M)\supset(\exists y)(має(x,y)\vee має(y,x))$ .

2.22. **Відносне** (В) – сутність у взаємовідношенні до деякої іншої сутності. Для неї повинно виконуватись відношення " $x$  має  $y$ ":  $(\forall x:B)\supset(\exists y)(має(x,y)\vee має(y,x))$ . Для любого відносного  $x$  повинен існувати деякий  $y$  такий, що  $x$  має  $y$  або  $y$  має  $x$ .

2.23. **Проміжне** (Пр) – сутність, яка структурує інші сутності у взаємозалежності.

2.31. **Тривале** (Т) – сутність, чия індивідуальність продовжує бути такою, що розпізнається на деякому розширеному інтервалі часу. Для неї задовольняють такі аксіоми:

- тривале  $x$  має тільки просторові частини і ніяких часових. В будь-який час  $t$ , коли  $x$  існує, всі із  $x$  існують в той же самий час  $t$ ;
- індивідуальні умови для тривалого є незалежними від часу.

Матеріальне тривале – об'єкт і абстрактне тривале – схема, яка може використовуватись для характеристики деякого об'єкта.

2.32. **Подійне** (Пд) – сутність, що не має стійкої ідентифікації на протязі любого інтервалу часу. Для неї задовольняють такі аксіоми:

- часові частини подійного, які називають стадіями, існують у різний час;
- просторові частини подійного, які називають учасниками, можуть існувати в той же самий час, але подійне може мати різних учасників на різних стадіях;
- немає ніяких умов ідентичності, які можуть використовуватись, щоб ідентифікувати дві події, що не перекриваються в області координат простір-час.

Третій і четвертий рівні ієрархії категорій створені перетином категорій більш високих рівнів.

Для третього рівня відведено такі категорії.

### 3.1. **Дійсне** (МН)

Матеріальна сутність (М), чиє існування є незалежним від будь-якої іншої сутності. Як елементи категорія "Дійсне" включає категорії "Об'єкт" і "Процес".

### 3.2. **Приєднання** (МВ)

Матеріальна сутність (М) зв'язана певним відношенням з деякою сутністю або сутностями. Для неї виконується перевірка відношення " $x$  приєднати  $y$ " з метою, чи є сутність  $y$  включеною в сутність  $x$ . Якщо так, то категорія "Приєднання" може бути виражена " $\text{приєднати}(x, y)$ ".

### 3.3. **Форма** (АН)

Абстрактна (А) інформація є незалежною від будь-якого кодування або варіанта реалізації. Форми можуть існувати в тому ж самому смислі, як і математичні об'єкти типу множин чи відношень, але елементи форм не можуть існувати в специфічному місці і часі без деякого матеріального кодування або варіанта реалізації. Їх ще називають "вічними об'єктами", тому що вони незалежні від простору.

### 3.4. **Судження** (АВ)

Абстракція (А) зв'язує певним відношенням деяку сутність або сутності. В логіці, твердження про судження – вимога, що абстракція відповідає деякому аспекту чи конфігурації сутності, чи залучених сутностей. Складні судження представляються конструкціями складеного предикату типу математичного виразу чи діаграми.

### 3.5. **Намір** (АПр)

Абстракція (А) розглядається як посередниця між інших сутностей. Четвертий рівень представлено такими категоріями.

#### 4.1.1. **Об'єкт** (МНТ)

Дійсність (МН) розглядають як тривале, що зберігає її тотожність більше деякого проміжку часу. Хоча не існує матеріальної сутності, яка завжди залишається постійною, об'єкт може бути розпізнаний відповідно з характеристиками, які залишаються стійкими впродовж їх життєвого циклу.

#### 4.1.2. **Процес** (МНПд)

Дійсність (МН) розглядають як подійне. В залежності від масштабу часу і рівня деталізації, одна й та сама дійсна сутність може розглядатися як стійкий об'єкт або динамічний процес. Навіть алмаз можна було б вважати процесом, якщо розглядався б довгий період часу чи на атомному рівні вібруючих частинок.

#### 4.2.1. **З'єднання** (МВТ)

Приєднання (МВ) розглядають як тривале більше деякого інтервалу часу. Сутність, що приєднується, є об'єкт у стійкому взаємовідношенні до деякої тривалої сутності, що приєднує. Приклади з'єднань включають вузол в ряді чи підключені частини автомобіля.

#### 4.2.1. **Участь** (МВПд)

Приєднання (МВ) розглядається як подійне впродовж інтервалу часу, що представляє зацікавленість. Сутність, яка приєднується, є процес у стійкому взаємовідношенні до деякої тривалої сутності, що приєднує.

#### 4.3.1. **Структура** (АНТ)

Форма (АН) має структуру тривалого. Схема – абстрактна форма, чия структура не визначає час або часоподібні взаємовідношення. Приклади включають геометричні форми, синтаксичні структури речень на деякій мові і кодування зображень в мультимедіа системі тощо.

#### 4.3.2. **Сценарій** (АНПд)

Форма (АН) має структуру події. Сценарій – абстрактна форма, яка представляє час чи часоподібні послідовності. Включає форми всього, що існує в постійному русі. За принципом дихотомії Сценарій може мати підтипи Процедура і Кінетична форма. Прикладами сценарію є комп'ютерна програма, рецепт для випічки пирога чи ноти для виконання музики, кінофільм тощо.

#### 4.4.1. **Опис** (АВТ)

Судження (АВ) відносно тривалості. Опис – судження, що формулює, як саме деяка схема характеризує деякий аспект (або конфігурацію) тривалого, матеріального чи абстрактного.

#### 4.4.2. **Хронологія** (АВПд)

Судження (АВ) відносно подійного. Хронологія – судження, яке зв'язує деякий сценарій зі стадіями деякого подійного. Комп'ютерна програма, наприклад, є сценарій; комп'ютер, що виконує програму, забезпечує процес, а абстрактна інформація, закодована в потоці команд, є хронологія. Як любе судження, хронологія може не бути істинною.

#### 4.5.1. **Причина** (АПрТ)

Причина є намір (АПр) деякого агента відносно деякого тривалого. На відміну від опису, причина пояснює сутність в термінах наміру.

#### 4.5.2. **Ціль** (АПрПд)

Ціль є намір (АПр) деякого агента, який визначає взаємодію сутностей в ситуації. Основні аксіоми для Цілі успадковані від "своїх" категорій вищих рівнів Абстрактного, Проміжного і Подійного. Вони зв'язують цілі з діями і агентами:

- Послідовність часу. Якщо агент  $x$  виконує дію  $y$ , чия ціль – місцезнаходження  $z$ , початок  $y$  передре початку  $z$ .
- Випадковість. Якщо агент  $x$  виконує дію  $y$ , чия ціль – місцезнаходження  $z$ , описане судженням  $p$ , то можливо, що  $z$  не міг би відбутися або що  $p$  не міг би бути істинним для  $z$ .
- Успіх або невдача. Якщо агент  $x$  виконує дію  $y$ , чия ціль – ситуація  $z$ , описана судженням  $p$ , то  $x$  є успішним, якщо  $z$  відбувається і  $p$  істинний для  $z$ ; інакше  $x$ , як кажуть, потерпів невдачу.

## 5. Висновки

В роботі представлено абстрактну модель мовно-онтологічної картини світу та ієрархію категорій верхнього рівня, згідно з якою повинно розгортатись онтологічне дерево сутностей в МОКС. Опираючись на ретроспективний аналіз досліджень онтологічних категорій [19], узагальнено принципи їх побудови, узагальнення і розділу (від простої дихотомії до комбінаційного методу Лейбніца).

Суттєве зростання рівня "інтелекту" сучасних інформаційних систем, зокрема, обумовлює розробку ефективних засобів інтерпретації як об'єктової, так і акторної складових семантики ПМТ, що, у свою чергу, потребує розробки єдиної методології побудови як загальної (мета) онтології, так і

онтологій з конкретних предметних областей, а також єдиної метамови формального представлення інформаційних одиниць синтаксису та семантики.

Крім того, ще чекають на розробку як формалізовані методики, так і комп'ютерне представлення задекларованих ациклічних орієнтованих графів для дієслова, прикметника та прислівника. Задача композиції вказаних графів постає проблемою, що передбачає окремий комплекс фундаментальних досліджень.

## СПИСОК ЛІТЕРАТУРИ

1. Андон Ф.И., Яшунин Л.Е., Резниченко В.И. Логические модели интеллектуальных информационных систем. – К.: Наукова думка, 1999. – 397 с.
2. Гладун В.П. Процессы формирования новых знаний. – София: СД "Педагог 6", 1994. – 192 с.
3. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. – 384 с.
4. Маторин С.И. Системологическое исследование структуры системы категорий // НТИ. –1997.– № 3. – С. 3–7.
5. Наталья Ф. Ной и Дэбора Л. Мак Гиннесс. Разработка онтологий 101: руководство по созданию Вашей первой онтологии. – [http://protege.stanford.edu/publications/ontology\\_development/ontology101.html](http://protege.stanford.edu/publications/ontology_development/ontology101.html).
6. Овдей О.М., Проскудина Г.Ю. Обзор инструментов инженерии онтологий. – <http://www.w3c.org/TR/1999/REC.html>.
7. Палагин А.В. Организация и функции "языковой" картины мира в смысловой интерпретации ЕЯ - сообщений // Information Theories and Application. – 2000. – Vol. 7, N 4. – С. 155–163.
8. Палагин А.В., Яковлев Ю.С. Системная интеграция средств компьютерной техники. – Винница: «УНІВЕРСУМ-Вінниця», 2005. – 680 с.
9. Широков В.А. Феноменологія лексикографічних систем. – К.: Наукова думка, 2004. – 327 с.
10. Соколовская Ж.П. «Картина мира» в значениях слов. – Симферополь: «Таврия», 1993.
11. Широков В.А. Інформаційна теорія лексикографічних систем. – К.: Довіра, 1998. – 331 с.
12. Апресян Ю.Д. и др. Лингвистический процессор для сложных информационных систем. – М.: Наука, 1992. – 287 с.
13. Кондаков Н.И. Логический словарь-справочник. – М.: Наука, 1975. – 720 с.
14. Dobrov B., Loukachevitch N., Nevzorova O. An Approach to New Ontologies Development: Main Ideas and Simulation Results. – International Journal "Information Theories & Applications". – 1998. – Vol.10.
15. Guarino N. Formal Ontology and Information Systems // N. Guarino (ed.) *Formal Ontology and Information Systems*. Proc. of FOIS'98. – Trento, Italy. – 1998. – 6–8 June.
16. Guarino N. Some Ontological Principles for Designing Upper Level Lexical Resources // Proc. of First International Conference on Language Resources and Evaluation. – Granada, Spain. – 28–30 May.
17. Guarino N. Concepts, Attributes and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases // *Data & Knowledge Engineering*. – 1992. – Vol. 8. – P. 249–261.
18. Экспертные системы для персональных компьютеров: методы, средства, реализации: Справочное пособие / Крисевич В.С., Кузьмич Л.А., Шиф А.М. и др. – Минск: Высшая школа, 1990. –197 с.
19. John F. Sowa, Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks Cole Publishing Co. – Pacific Grove, CA, 2000.