

УДК 004.5:004.6:004.89:007.51:528.933

**В.В. ПРИХОДНЮК, О.Є. СТРИЖАК**

## **МНОЖИННІ ХАРАКТЕРИСТИКИ ОНТОЛОГІЧНИХ СИСТЕМ**

***Анотація.** Розглядається множинність бінарних відношень між концептами онтологічних систем. Визначаються особливості утворень таксономічних структур на основі використання гіпервідношення множинної впорядкованості. Введені та визначені властивості множинної впорядкованості. Визначено їх вплив на процеси формування різноманітних таксономічних структур і категорій. Формулюються умови щодо визначення істинності множинного гіпервідношення, яке виражає властивість індуктивності впорядкованих множин концептів таксономій. Описуються функціональні властивості множинних характеристик онтологічних систем.*

***Ключові слова:** онтологія, таксономія, множинність, бінарне відношення, впорядкованість, гіпервідношення, властивість.*

### **Вступ**

В мережевому інформаційному просторі (Information networks space) (далі – МІП) [1, 2] постійно утворюються, накопичуються і архівуються великі об'єми неструктурованої і слабо структурованої інформації. Актуальність цих інформаційних ресурсів дуже велика, так як в них надаються знання і технологічні рішення, які відображають розвиток сучасної науки і техніки. Також вони зберігають в собі результати спостережень, моніторингу і аналізу різних процесів, які реалізуються в суспільстві. Багатоаспектність тематик такої інформації призводить до проблеми її некоректної інтерпретації, особливо, коли необхідно забезпечити в середовищі МІП взаємодію між колективами спеціалістів з різних Про. Зазвичай, необхідність в такій взаємодії виникає при вирішенні складних проблем, які розбиваються на не менш складні предметні прикладні задачі.

Характер використання мережевих інформаційних ресурсів, в процесі вирішення різноманітних прикладних задач, визначається інструментами (програмними засобами) отримання, обробки та зберігання, дослідженням та реалізацією яких займається така наукова дисципліна, як інформатика [3]. Один з напрямків таких досліджень пов'язаний з процесами формалізації концептуального відображення описів предметних областей, визначається на основі принципів методології онтологічного моделювання пасивних і активних інформаційних процесів накопичення, обробки, відображення та взаємодії [4, 5, 6].

Використання онтологічних систем для відображення семантики інформаційних ресурсів МІП [7, 8, 9] у вигляді ієрархічних структурних специфікацій [10, 11, 12, 13], над якими задається певна розширювана аксіоматика і між якими визначаються множини відношень, дозволяє частково вирішити проблему коректної інтерпретації їх використання при вирішенні складних прикладних задач.

## 1. Структурування на основі множинної впорядкованості

Основу будь-якої онтологічної системи  $O$  складає певна структурна специфікація [4, 5, 7, 11], однією з властивостей якої є відображення відношень, що існують між термінами-концептами, які складають та визначають її предметну тематику. Зазвичай основу структурування складають класи об'єктів [14], на які умовно розбивається множина всіх термінів  $X = \{x_i / i = \overline{1, n}\}$ , властивості яких визначають їх тематичні семантики. Безпосередньо властивості концептів дозволяють визначити вказану вище множину відношень між ними –  $R$ . Над цією множиною може бути задане гіпервідношення бінарної впорядкованості  $Gr$ , яка має властивості ациклічності –  $\gamma$ , часткової впорядкованості –  $\tilde{p}$  та лінійної впорядкованості –  $p$  [10, 15–18].

Вказані властивості зв'язані між собою наступними правилами перетворень [17, 18]:

$$x_i^j \gamma x_l^k \xrightarrow{\alpha} x_i^j p x_l^k \quad (1)$$

$$x_i^j \tilde{p} x_l^k \xrightarrow{\alpha} x_i^j p x_l^k \quad (2)$$

Бінарні відношення ациклічності і часткового порядку дозволяють формувати із загальних концептів предметної області множину таксономій  $\tilde{T}$  [11, 15], яка й визначає поверхневу структурну специфікацію онтологічної системи. Більш того, згідно з правилами (1) і (2), на основі множини усіх бінарних пар концептів таксономічної структури, можуть бути утворені певні тавтології [16], які визначають істинні твердження щодо тематичних властивостей концептів.

Розгляд множинних відношень виду (1) і (2), утворених між концептами, включаючи множини всіх множин концептів виду  $\{x_i\}$  онтологічної системи  $O$ , і самими множинами, дозволяє їх типізувати як множинні бінарні відношення часткової впорядкованості –  $\tilde{P} \subset Gr$ . Згідно з правилом (2),  $\tilde{P}$  може бути перетворено до множинного бінарного відношення строгої часткової впорядкованості  $P$ , яка дозволяє визначити між концептами онтологічної системи бінарне відношення лінійного порядку  $p$ . Тобто можна завжди для бінарного відношення часткового порядку, яке визначається правилом виду (1), знайти монотонно зростаючу послідовність з множин натуральних чисел  $N$ , що визначають  $n$ -й порядок бінарної взаємодії концептів онтологічної системи. Така бінарна гіпервпорядкованість може бути визначена у вигляді множини бінарних лінійних впорядкованостей  $P$  і представлена у вигляді наступного виразу –  $P = \{p_n\} n \in N$ :

$$x \gamma y \Leftrightarrow \forall n \in N : x p_n y \quad (3)$$

Розглядаючи множини гіпервідношень  $Gr$ , можна виділити наступні властивості, що характеризують їх як бінарні [16, 17, 18, 19, 20]:

– агіперциклічність – якщо для  $Gr$  не існує гіперциклічної множини  $X \subseteq U$  такої, коли:

$$\forall x \in X \exists Y \subseteq X : YGrx ; \quad (4)$$

– іррефлексивність:

$$YGrx \Rightarrow (Y / \{x\})Grx ; \quad (5)$$

– гіпертранзитивність:

$$YGrx, x \in X, XGrz \Rightarrow ((Y \cup X) / \{x\})Grz ; \quad (6)$$

– регулярність:

$$YGrx, Y' \supseteq Y \Rightarrow Y'Grx . \quad (7)$$

Властивість множинної упорядкованості таксономічних категорій дозволяє формувати та/або виділяти з множин концептів онтологічної системи підмножину концептів, пов'язаних між собою конкретним бінарним відношенням впорядкованості. Таким чином, якщо множина гіпервідношень володіє набором властивостей (3)–(7), забезпечує відображення різних активних станів онтологічної системи  $O$  на основі виділення бінарних відношень часткової впорядкованості  $\tilde{p}$  для множин концептів  $\{x_i\}$ , то завжди можна визначити таксономію та/або таксономічну категорію  $T \in \tilde{T}$ , для концептів якої можуть бути сформульовані предикативні вирази, що приймають значення істинності. Тобто ми завжди можемо з множини концептів  $X$  виділити непорожню підмножину концептів  $\{x\}$ , що має властивість бінарного відношення часткової впорядкованості  $\tilde{p}$ , яке і пов'язує їх між собою:

$$YGrx \Rightarrow \exists y \in Y : y\tilde{p}x . \quad (8)$$

При цьому часткова впорядкованість є елементом множини бінарних впорядкованостей –  $P$ , яка може бути розширена до гіпервідношення  $Gr$  й тим самим визначати множинну впорядкованість над підмножинами концептів  $\{x\}$  у вигляді множинного відношення – «група об'єктів – об'єкт», що має властивості (4)–(8):

$$YGrx \Leftrightarrow n \in N \exists y \in Y : yp_n x : p_n \in P \subseteq Gr . \quad (9)$$

Предикативні вирази, що можуть бути сформульовані на основі концептів таксономічної категорії із заданим множинним відношенням впорядкованості виду (1)–(9), приймають лише значення істинності. Це дозволяє формувати на основі термінів-концептів онтологічної системи лінгвістичні вирази, що осмислено відображають стан онтологічної системи.

Тоді довільні концепти з множини  $X = \{x_i / i = \overline{1, n}\}$  можуть бути впорядковані у вигляді:

$$r^m (x_i^j, x_l^k \mid x_i^j \in X_i; x_l^k \in X_l; r^m \in R \neq \emptyset), \quad (10)$$

де  $r$  – певне бінарне відношення між концептами множини  $X = \{x_i / i = \overline{1, n}\}$ , з множини відношень  $R$ ,  $r \in R$ .

Тобто завжди є певна непорожня множина істинних висловлювань, яка утворюється концептами множини таксономії  $\tilde{T}$  онтологічної системи, за умови, що ці концепти впорядковані між собою бінарними відношеннями виду (10).

Так, для множини таксономічних категорій  $\tilde{T} = \{\text{тип (phylum); підтип (subphylum); клас (classis); підклас (subclassis); ряд (у рослин – порядок) (ordo); підряд (subordo); родина (familia); підродина (subfamilia); рід (genus); підрід (subgenus); вид (species); підвид (subspecies); різновидність (varietas); форма (forma)}\}$ , множина бінарного відношення часткової впорядкованості  $P$  забезпечує збереження всіх типів взаємодії між ними. Це також дозволяє формувати всі множини таксономії  $\tilde{T}$  з концептів сформованої онтологічної системи. Однак задавши над зазначеними категоріями відношення лінійної впорядкованості, ми звужуємо їх перелік, тому що ряд категорій, таких як: клас (classis); форма (forma); вид (species) можуть займати у виразі (2), що представляє бінарне відношення – «бути елементом категорії», як ліву, так і праву частину. Тим самим стають нездійсненними властивості відношення (5) і (7) – іррефлексивності та регулярності.

Тут однак слід відзначити наступну властивість гіпервідношення бінарної впорядкованості  $Gr$ : правила (1)–(9) повністю забезпечують формування з класів концептів  $\{x\}$  різних таксономічних структур множини  $\tilde{T}$ .

Слід зазначити, що таксономічна категорія формується на основі виділення певної підмножини концептів, що мають загальну властивість, яка їх усіх характеризує. Така властивість може бути унарною для кожного концепту, проте при визначенні зазначеного класу як складного концепту, така властивість дозволяє визначити над усіма концептами класу множинне бінарне відношення «бути елементом класу». При цьому слід зазначити, що унарна властивість, що є загальним для множини концептів предметної області, може інтерпретуватися як ознака цих концептів або як критерій для вибору концептів даного класу.

Так, під час вибору концептів-кандидатів для формування певної категорії – клас (classis), спочатку визначається властивість-ознака, на підставі якої дана категорія може бути представлена як складний концепт. При цьому включення до категорії-класу конкретного простого концепту або поняття

з менш складною структурою ґрунтується на унарній *властивості-критерії*, що характеризує його та є загальною для всіх концептів цієї *категорії-класу*. Однак вказана загальна властивість дозволяє визначити для даної категорії бінарне відношення «*частина-ціле*», яке також задає часткову впорядкованість виду (2) над усіма концептами сформованої таксономічної категорії.

Таким чином, множина бінарного відношення часткової упорядкованості може бути задана на основі бінарного відношення «*частина-ціле*», яке можна переписати у вигляді множинного відношення – «*група об'єктів – об'єкт*». Тут фраза «*група об'єктів*» визначає ім'я складного *концепту-класу*, який є таксономічною категорією для розглянутої предметної області. Вибір концептів для їх включення в таксономічну категорію за загальними семантичними властивостями ми можемо розглядати як об'єднання одноелементних множин, де кожна така множина визначена конкретним концептом предметної області з обраною загальною для них *властивістю-критерієм*.

Конструктивність застосування процедури вибору тут проявляється можливістю точного визначення необхідного елемента для формування таксономії. Перевага критеріального вибору в даному випадку не залежить від методу ранжування. Таксономія задається концептами із загальною властивістю, над якими визначається множинне бінарне відношення часткової впорядкованості виду (2).

Таксономія може бути визначена для будь-якого складного концепту, тобто за замовчуванням вона може бути утворена бінарним відношенням лінійної впорядкованості «*частина-ціле*». Відношення «*частина-ціле*» може бути розширено до множинного відношення «*група об'єктів – об'єкт*» і далі до «*бути елементом класу*» та/або «*бути елементом категорії*». Відмінність між поняттями «*клас*» і «*категорія*» полягає в повноті відображення тематики предметної області. Поняття «*категорія*» досить повно відображає семантичні властивості предметної області, водночас поняття «*клас*» визначається просто відібраною множиною концептів із загальними семантичними властивостями. Тому поняття «*таксономія*»  $T$  не може бути повним еквівалентом поняття «*таксономічна категорія*»  $\tilde{T}$ .

Гіпервідношення  $Gr$ , задане над множинними бінарними відношеннями впорядкованості виду (1) і (2), що визначають таксономію, дозволяє уявити предикативні вирази, що формулюються на основі її концептів із заданим множинним відношенням упорядкованості виду (1)–(9) у вигляді рекурсивного предиката наступного виду [13, 16, 17]:

$$Pr(x_1, \dots, x_n) = \begin{cases} 1, \neg Pr(YGrx_i) \wedge Pr(x_1, \dots, x_n) \\ 0, Pr(YGrx_i) \end{cases} \quad (11)$$

де  $x_i \in X; 1 \leq i \leq n$ .

Фактично предикат виду (11) дозволяє визначити, які концепти онтологічної системи володіють даною властивістю. Сформувавши множину класів концептів онтологічної системи за допомогою предикативного виразу

виду (11), ми отримуємо таксономію, над концептами якої задано множинне бінарне відношення «*частина-ціле*». Така таксономія може мати складну структуру низхідних ієрархій. Кожний складник являє собою клас концептів, що мають як мінімум одну загальну семантичну властивість.

$$Pr(x_1, \dots, x_n) = 0 \Rightarrow \exists T \subseteq \tilde{T} : \forall x \in X \exists Y \subseteq X \neg T \ YGrx. \quad (12)$$

Гіпервідношення *Gr* виду (1)–(9) визначають взаємодії між концептами кожної таксономії, виділеної з різних класів концептів предметної області. Відповідно рекурсивний предикат виду (11) дозволяє визначити нові види таксономії онтологічної системи. Особливу роль у формуванні нових таксономій відіграє, крім рекурсивних предикатів, функція вибору. У термінах таксономічних категорій вона має наступний вигляд:

$$\forall T \left[ \emptyset \notin T \Rightarrow \exists F : T \rightarrow \cup T, \forall T \in \tilde{T} (F(T) \in \tilde{T}) \right]. \quad (13)$$

Тоді характеристична функція рекурсивного предиката (11) може визначати умови застосовності функції вибору при формуванні множини таксономій  $\tilde{T}$  онтологічної системи *O*. Якщо значення характеристичної функції дорівнює нулю, то ми виділили множину концептів, з яких неможливо сформувати істинні висловлювання, тому вони не володіють жодним з бінарних гіпервідношень впорядкованості *Gr*. В протилежному випадку характеристична функція приймає значення 1, а з множини обраних концептів формулюються істинні висловлювання.

Відповідно, над ними можуть бути визначені гіпервідношення *Gr*. Тобто завжди можна знайти непорожню множину концептів онтологічної системи, де існує хоча б один з типів відношень впорядкованості виду (1)–(7), і концепти, пов'язані таким відношенням, можуть утворювати істинні висловлювання. Таким чином, ця умова задає індуктивність процесу формування множин концептів, між якими встановлюються гіпервідношення *Gr*, і фактично визначає правила конструювання таксономій та/або таксономічних категорій. На підставі тверджень, наданих у вигляді (11)–(13), ця умова може бути представлена в наступному вигляді:

$$\exists P \subseteq Gr, \exists p \in P, \forall x \in X \exists Y \subseteq X : YGrx \Rightarrow \exists T = YGrx. \quad (14)$$

Індуктивність виразу (13) задається послідовністю застосування предикативних виразів виду (11) і (12) до множини концептів  $X = \{x_i / i = \overline{1, n}\}$  онтологічної системи *O*.

## 2. Властивості впорядкованості таксономій

Формування таксономічних структур на основі процедури вибору і множини бінарних гіпервідношень впорядкованості дозволяє визначити низку властивостей, що забезпечують конструктивність застосування таксономій  $\tilde{T}$

для реалізації різних станів онтологічної системи  $O$ . Особливістю нашого підходу є алгебраїчне представлення вибору відповідної переваги при формуванні множин класів з концептів предметної області.

Індуктивність предикативної форми представлення процесу формування таксономічних структур (11) і (12) розширена умовою існування таксономії (14), що дозволяє визначити ряд конструктивних властивостей таксономій як гіпервпорядкованих бінарних структур. Зазначена індуктивність накладає низку умов на вхідні і вихідні дані [4, 15] онтологічної системи при виборі концептів, на основі яких формується таксономія. Якщо вхідні і вихідні дані онтологічної системи  $O$  визначають стійкість формування таксономічних структур, то регулювати їх взаємодію на підсистемному рівні можливе при дотриманні ряду умов.

Ці умови визначаються типом відношень бінарної впорядкованості (1)–(7) і забезпечують незалежність значень характеристичної функції рекурсивного предиката (11) від послідовності контекстів індуктивного вибору. У складних системах, якими є онтології, виділяють наступні умови стійкості [10, 19, 20]:

Умова спадкування:

$$\exists T' \subseteq T \Rightarrow \exists F(T') \supseteq F(T) \cap T' . \quad (15)$$

Умова незалежності від концептів, які не входять до:

$$\exists F(T') \subseteq T' \subseteq T \Rightarrow F(T') = F(T)' . \quad (16)$$

Умова згоди:

$$\bigcap_n F(T_n) \subseteq F(\bigcup_n T_n) . \quad (17)$$

Властивості таксономічних структур, представлені виразами (15)–(17), дозволяють певною мірою проінтерпретувати інтуїтивні міркування при виборі конкретних властивостей концептів, на підставі яких реалізується вибір переваг щодо їх включення до таксономії  $T$ . Так, умова успадкування забезпечує включення концептів, що мають еквівалентні *властивості-критерії*. Умова незалежності дозволяє задати набір аксіом для онтології, що виключають виникнення протиріч при виборі. Умову згоди дозволяє забезпечити коректність розбиття множини концептів предметної області онтологічної системи на класи. При цьому згідно з дослідженнями [10, 13, 19, 20] забезпечується формування класу функцій, що визначають достатньою мірою множинність бінарних відношень упорядкованості. Більше того, можна стверджувати, що умови (15)–(17) досить точно визначають саму систему вибору за умови, що *властивості-критерії*, на основі аналізу яких реалізується формування таксономії, задані.

Представлення властивостей концептів таксономії може бути реалізоване у вигляді предикативного вираження (12). Тоді застосування умов (15)–(17) до концептів предметної області онтології дозволяє формувати таксономічні структури – класи та/або категорії у вигляді дерев ознак [20]. Це забезпечує можливість представлення таксономій, концепти яких взаємодіють один

з одним на основі множини гіпервідношень  $\{Gr\}$ , у вигляді графа без циклів [4, 15].

На основі множинної гіпервпорядкованості таксономій завжди виконується наступне твердження: якщо певна онтологічна система  $O_n$  формується на основі таксономії  $T$ , то завжди є певна непорожня множина істинних висловлювань  $\{Pr(x_1, \dots, x_n)\}$ , яка утворюється концептами  $x$  цієї онтології, за умови, що ці концепти пов'язані між собою бінарними відношеннями виду (15)–(17).

Істинність цього твердження виводимо з виразу (12), що можна проінтерпретувати наступним чином: будь-яка ієрархічна структура, для якої є справедливими умови успадкування, незалежності та згоди (15)–(17), може бути представлена у вигляді бінарного дерева і, таким чином, у вигляді певної таксономії. Це твердження носить конструктивний характер, тому що забезпечує індуктивне формування правил обробки концептів з різних таксономій на основі виявлення між ними бінарних відношень множинної впорядкованості.

Додатково слід зазначити: умова згоди (17) визначає властивість пластичності множини таксономій  $\tilde{T}$  онтологічної системи  $O$ . Властивість пластичності дозволяє стверджувати, що для довільної онтологічної системи завжди існує непорожня множина тавтологій, яка може бути розширена новою тавтологією. Таким чином, при використанні певних концептів онтології можливо виконувати підстановку множин висловлювань, які задаються на множині концептів онтології на підставі їх тематичних властивостей і бінарних відношень впорядкованості виду (1)–(7). При цьому всі зазначені концепти мають бути представлені у вигляді предикативного виразу (12). Це твердження може бути перефразовано таким чином: якщо певна онтологія  $O$  формується на основі множини таксономій  $\tilde{T}$ , множини концептів  $X$ , то завжди є певна непорожня множина висловлювань  $\{Pr(x_1, \dots, x_n)\}$ , яка утворюється концептами  $x$  цієї онтології, за умови, що ці концепти пов'язані між собою бінарними відношеннями виду (1)–(9). Таким чином можна визначити відношення множинної гіпервпорядкованості у вигляді наступного виразу:

$$\exists X, \exists T \Rightarrow X Gr T. \quad (18)$$

Вираз (18) можливо проінтерпретувати наступним чином. В довільній онтології  $O$  завжди існують концепти  $x_i$ , які мають множинні бінарні гіпервідношення  $Gr$  між собою.

### 3. Безтипове представлення функціональності множинної впорядкованості



Згідно з визначенням онтологічної системи  $O$  [4–9, 21–23, 28, 29], одним з її утворюючих системних компонентів є множина інтерпретуючих функцій –  $F$ . Фактично ці функції визначають правила щодо використання та обробки концептів онтології  $X = \{x_i / i = \overline{1, n}\}$  при вирішенні певних задач та проблем. Застосування до множини концептів  $X$  онтологічної системи  $O$ , над якими задані бінарні відношення типу (1), (2), (10), будь-якого погодженого правила з множини інтерпретуючих функцій  $F$ , які можуть бути представлені у вигляді функції типу  $f^i(x_1, \dots, x_n, x_{n+1}) \in F_k$ , дозволяє для довільної онтологічної системи  $O$  завжди визначити непорожню множину тавтологій, що може бути розширена новою тавтологією.

Таким чином, при використанні онтологій в процесі формування лінійно-впорядкованих таксономій, можна виконувати підстановку множин висловлювань, які задаються на множині концептів  $X = \{x_i / i = \overline{1, n}\}$  у вигляді тавтологій. Вказані множини висловлювань та тверджень складають певні текстові масиви, утворюючи множини термінів-концептів яких –  $X$ , пов'язані між собою різними множинами семантичних відношень  $R_{sem} \subseteq R$ . Як вже було відзначено, над відношеннями з кожної такої множини  $R_{sem}$  може бути задане відношення часткової впорядкованості  $\tilde{p}$ . Це відношення дозволяє формулювати множини правил  $Rul$ , на підставі застосування яких ми можемо формулювати істинні твердження з концептів ПрО. При цьому будемо вважати, що концепт є вільним, якщо не пов'язаний з іншими концептами з ПрО ніякими типами відношень. Якщо ж між концептами встановлені які-небудь відношення з множини  $R_{sem}$ , то тоді ми такі концепти будемо визначати як пов'язані. Тоді усі висловлювання та твердження утворюються пов'язаними концептами. А коректність визначення над ними множинного відношення часткової упорядкованості  $\tilde{p}$  дозволяє вважати їх виконуваними або істинними [16, 24].

Побудова тверджень з концептів реалізується на підставі кінцевих наборів правил з множини  $Rul$ , які визначають порядок застосування множинного відношення часткової упорядкованості  $\tilde{p}$  як над концептами з множини  $X$ , так і над семантичними відношеннями з множини  $R_{sem}$ . Також відношення  $\tilde{p}$  дозволяє нам формувати з концептів множини  $X = \{x_i / i = \overline{1, n}\}$  відповідні термінополя [6, 13, 25, 26] у таксономій з множини  $\tilde{T}$ , де між концептами задається множинне бінарне відношення часткової упорядкованості –  $\tilde{p}$ .

Такі правила побудови тверджень носять аплікативну форму і можуть бути представлені у вигляді безтипового виразу виду [15, 27]:

$$f_a = (\lambda x.t(x))a = t(a), \quad (19)$$

де:  $\lambda$ -теорія – лямбда числення; запис  $\lambda x$  говорить, що це  $\lambda$ -терм;

$x$  – змінна, що приймає значення на множині лексем  $L$  або концептів  $X$ , де  $L = \langle x_i | i = \overline{1, n} \rangle$ ,  $x \in X = \{x_i | i = \overline{1, n}\}$ ;

$t$  – вираз, який може містити змінну  $x \in X = \{x_i | i = \overline{1, n}\}$ ;

$a$  – аргумент функції (контекст), що визначає можливі значення змінної  $x$ , такої що,  $x \in X = \{x_i | i = \overline{1, n}\}$ ;

$f_a$  – функція, що може бути застосована до аргументу  $a$ .

Як можемо бачити з виразу (19), змінні  $x$  дозволяють визначити властивість екстенціональності множини концептів, що також є фінітним [6, 10, 27], так як фактично екстенціональність концепту як певного терма  $\lambda$ -теорії визначається його приналежністю до певного класу. Тоді ми завжди можемо обмежити істинність твердження, ввівши аксіому, що виключає істинність твердження для концептів, які входять в один клас.

Перевагою даного аплікативного виразу є те, що в якості змінної і аргументу, окрім концептів множини  $X$ , можуть бути відношення множини  $R$ . Тобто, ми можемо будувати правила, які пов'язують між собою і концепти, і відношення. Тоді правила виду (19), що включають в себе в якості змінних і аргументів відношення з множини  $R_{sem}$  і множинне відношення  $\tilde{P}$ , мають властивість до інтенціональності. Тобто твердження, які мають властивість істинності, володіють ще набором інших властивостей, що визначають придатність даного твердження.

Функціональність множинної впорядкованості концептів та їх класів з утворюючої множини  $X = \{x_i | i = \overline{1, n}\}$  представима застосуванням правил з множини  $F$  до множини таксономій  $\tilde{T}$ , що досягається на основі виконання наступних етапів їх обробки.

Ідентифікація множини термінів-концептів  $X$ , що належать заданому термінополю множини  $X$ .

Ідентифікація множини семантичних відношень  $R_{sem}$  між концептами.

Ідентифікація множини атрибутів  $A$  (контекстів) концептів множини  $X$ .

Вказані етапи ідентифікації можливі за умови виконання послідовного перетворення вхідної множини лексем (висловлювань та тверджень) –  $L$  за допомогою послідовного застосування правил з множини  $Rul_{\subseteq} F$ .

На множині лексем  $l \in L$  за допомогою оператора " $<$ " (*передус*) визначено лінійний порядок, і, таким чином,  $L$  є лінійно впорядкованою множиною:  $l_1 < l_2 < \dots < l_n$ . Також лексеми розбиті по реченнях  $S_i$ , на множині яких аналогічним чином задано лінійний порядок:  $S = \{S_1 < S_2 < \dots < S_m\}$ .

Кожне висловлювання також являє собою лінійно впорядковану множини:  $S_i = \{l_1^i < l_2^i < \dots < l_{n_i}^i\}$  й має аплікативний характер та може бути представлено у вигляді виразів виду (19). Правила застосовуються окремо до кожного з висловлювань –  $S_i$  і діють виключно в рамках його контексту  $a \in A$ .

Кожне таке правило задає перетворення одного з видів (2)–(4):

$$L \xrightarrow{Rul} X \quad (20)$$

$$X \xrightarrow{Rul} \langle X, R_{sem} \rangle \quad (21)$$

$$L \xrightarrow{Rul} A. \quad (22)$$

Крім того, можливі допоміжні перетворення:

$$L \xrightarrow{Rul} L \quad (23)$$

$$L \xrightarrow{Rul} L^* . \quad (24)$$

Де множина  $L^*$  – множина конструктів. Конструкт об'єднує в собі кілька лексем, які в подальшому обробляються як одна. Конструкти можуть мати такі ж самі характеристики, як і лексеми, і аналогічно до лексем можуть бути зв'язані з іншими лексемами чи конструктами синтаксичними зв'язками  $R_{syn}$ .

Будь-яке правило виду (20)–(4), а в деяких випадках – і виду (23) може бути застосоване не тільки до множини  $L$ , а і до множин  $L^*$  або  $L \cup L^*$ .

Першим і найбільш очевидним джерелом структури довільного висловлювання є його контекст (зміст). Контекст являє собою набір речень  $S_{toc} \subset S$ , які визначають відповідні концепти з множини  $X = \{x_i | i = \overline{1, n}\}$  й певним чином виділені з основного тексту висловлювання. Найчастіше під контекст відводиться кілька рядків. Тоді його достатньо легко ідентифікувати, задавши його межі і скориставшись гіпервідношенням впорядкованості  $Gr$  на кортежі  $L = \langle x_i | i = \overline{1, n} \rangle$ , елементи якого належать множині концептів  $X = \{x_i | i = \overline{1, n}\}$ .

Все, що необхідно в подальшому, це ідентифікувати місця у висловлюванні, на які посилаються елементи контекстів концептів термінополя. Це правило представимо у вигляді наступного виразу:

$$T = \lambda l_1, l_2 .. l_n, t, \quad (25)$$

$$t \equiv \exists i, \forall j \in [1, n_i], S_i \in S_{toc} \cup l_j^i \in S_i. \quad (26)$$

При відсутності контексту, необхідно сформулювати предикат  $q$  для аналізу розмітки і замінити умову (26) на (27):

$$t \equiv \exists i, \forall j \in [1, n_i], q(l_j^i). \quad (27)$$

Після застосування предикату ідентифікації виділені ним послідовності лексем формують множину категорій  $\{X_{cat}\}$ . Завдяки лінійному порядку лексем і контекстів можна розбити оригінальний текст висловлювання на частини:

$$L_i^{cat} \equiv \{l | \forall l^{i-1} \in S_{i-1}^{toc}, \forall l^{i+1} \in S_{i+1}^{toc}, l^{i-1} < l < l^{i+1}, \quad (28)$$

Кожну з множин  $L_i^{cat}$  можна обробляти як окреме висловлювання.

Категорії  $\{X_{cat}\}$  формують головні (верхні) категорії таксономії: всі виділені з фрагменту тексту висловлювання  $L_i^{cat}$  категорії являються підкатегоріями відповідної категорії  $X_i^{cat}$ .

Функціональне виділення концептів і зв'язків є досить складним процесом через велику варіативність лексичних конструкцій, що можливі на основі обробки контекстів кожного концепту з тексту висловлювання.

Для цього треба визначити спеціальну функцію Аналізатор –  $f_{an} \in \text{Rul} \subseteq F$ , для якої необхідно визначити правила формування формального опису таких конструкцій, і якість аналізу напряму залежить від повноти цього опису.

Опис представлений у вигляді правил виду (19). Конкретний вигляд правил залежить від типу правила і входної підмножини лексем, для обробки якої призначене це правило. Складовими правилами є предикати ідентифікації виду (29) і (30), які призначені для обробки кожної окремої лексеми

$$c_{a,b} = (\lambda x, y. t(x, y)) a, b \equiv \langle a, b \rangle \in LP, \quad (29)$$

$$r_{a,b,c} = (\lambda x, y, z. t(x, y, z)) a, b, c \equiv \langle a, b, c \rangle \in LS. \quad (30)$$

Для кожного предикату певним чином формується множина  $LP$  або  $LS$ .  $LP$  являє собою множину лексем і може бути визначена двома способами: простим переліком допустимих лексем або визначенням певної ознаки, що формує категорію таких лексем.  $LS$  являє собою множину пар лексем, що зв'язані певним видом синтаксичного зв'язку. Таким чином, кожен предикат визначається відповідною йому множиною допустимих лексем.

На основі таких предикатів формується правило виду (31):

$$rul = c_{x_1 p_1} \wedge c_{x_2 p_2} \wedge \dots \wedge c_{x_n p_n} \wedge r_{x_1 x_2 k_{12}} \wedge r_{x_2 x_3 k_{23}} \wedge \dots \wedge r_{x_{n-1} x_n k_{n-1n}}. \quad (31)$$

Застосування правила полягає у знаходженні впорядкованої множини лексем (14), для якої виконується умова (15).

$$L_{rul} \subset L, l_1^{rul} < l_2^{rul} < \dots < l_n^{rul}, \quad (32)$$

$$(\lambda x_1, x_2 \dots x_n. rul) l_1^{rul}, l_2^{rul} \dots l_n^{rul}. \quad (33)$$

Правила виду (20), (22)–(24) в подальшому виконують перетворення (34)–(38) відповідно:

$$\{l_1 \dots l_k, l_1^{rul} \dots l_n^{rul}, l_{k+n} \dots l_m\} \xrightarrow{rul} \{l_1 \dots l_k, x, l_{k+n} \dots l_m\} \quad (35)$$

$$\{l_1 \dots l_k, l_1^{rul} \dots l_n^{rul}, l_{k+n} \dots l_m\} \xrightarrow{rul} \{l_1 \dots l_k, a, l_{k+n} \dots l_m\} \quad (36)$$

$$\{l_1 \dots l_k, l_1^{rul} \dots l_n^{rul}, l_{k+n} \dots l_m\} \xrightarrow{rul} \{l_1 \dots l_k, l, l_{k+n} \dots l_m\} \quad (37)$$

$$\{l_1 \dots l_k, l_1^{rul} \dots l_n^{rul}, l_{k+n} \dots l_m\} \xrightarrow{rul} \{l_1 \dots l_k, l^*, l_{k+n} \dots l_m\}. \quad (38)$$

Правила формату (21) мають іншу структуру і виконують перетворення (39):

$$\{l_1^{rul}, l_2, \dots, l_{n-1}, l_2^{rul}\}^{rul} \rightarrow \langle \{l_1^{rul}, l_2, \dots, l_{n-1}, l_2^{rul}\}, \{ \langle l_1^{rul}, l_n^{rul}, R_{sem} \rangle \} \rangle. \quad (39)$$

Виділення кандидатів в атрибути відбувається під час виділення концептів правилами виду (22). В результаті їх застосування у вигляді перетворень виду (36), формується множина контекстів  $A^*$ . Для формування множини власне атрибутів (контекстів)  $A$  необхідно здійснити процедуру валідації елементів  $a \in A^*$  і відкинути ті з них, які не пройдуть валідацію.

Для кожного з можливих типів атрибутів створюється окремий предикат валідації  $q$ , що і визначає, чи повинна лексема входити до результуючої множини атрибутів. Предикати валідації можуть бути представлені у вигляді виразу виду (11), що визначає їх залежність від багатьох чинників, зокрема, типу контексту, підмножини мови опису контексту, що обробляється, предметної області. Наприклад, для географічних координат умовою валідності може бути належність координат певній робочій області, утворююче предикат-правило матиме вигляд (40), а сам предикат виду (11) представимо у вигляді виразу (41):

$$f_q = x_{min} < a_x < x_{max} \wedge y_{min} < a_y < y_{max} \quad (40)$$

$$q(a_x, a_y) = \begin{cases} 1, f_q(a_x, a_y) \\ 0, \neg f_q(a_x, a_y) \end{cases} \quad (41)$$

На основі таких предикатів формуються множини атрибутів (контекстів) по типах (42) і результуюча множина контекстів (43):

$$A^i = \{a^i | q_i(a^i)\} \quad (42)$$

$$A = \cup_i A^i. \quad (43)$$

Сформована множина контекстів-атрибутів  $A$  вигляду (42) і (43), характеризується тим, що над усіма його елементами задається гіпервідношення множинної впорядкованості  $Gr$ . Тоді на підставі застосування предикативного вираження виду (41), в ньому завжди можна виділити непорожню множину концептів, що утворюють бінарні пари виду:

$$\lambda((x_i)_{rul}) Gr \lambda((y_j)_{rul}), \quad (44)$$

кожен терм якого представляє певну лексему оброблюваних контекстів певного висловлювання з концептів множини  $X = \{x_i | i = \overline{1, n}\}$ . Конструктивною особливістю виразу (44) є можливість представлення кожної бінарної пари у вигляді тематичної тавтології [15].

Процес побудови множини таксономій  $\tilde{T}$  з концептів термінополя тексту висловлювань на основі застосування правил (1)–(43) може бути представлений у вигляді такої продукції:

$$\lambda((x_i)rul) Gr \lambda(y_j)rul \Rightarrow \tilde{T} = \lambda(x.t(x), Gr, <). \quad (45)$$

Правило (45) задає індуктивність процесу формування впорядкованих множин концептів виду (42), між елементами яких встановлюються гіпервідношення множинної впорядкованості, і фактично конструюється таксономія. Необхідною умовою індуктивності є визначення над концептами текстів предикативного вираження (40). Предикативні висловлювання формулюються на основі концептів таксономічної категорії із заданим гіпервідношенням множинної впорядкованості і приймають тільки значення істинності. Це дозволяє формувати на основі термінів концептів таксономічної системи лінгвістичні вирази, які змістовно відображають смислові стани висловлювань як текстових масивів, у вигляді пасивної системи тематичних знань.

## Висновки

Множинні характеристики, які визначаються на основі використання гіпервідношення множинної впорядкованості концептів тематичних предметних областей, забезпечують досить конструктивне визначення та використання властивості функціональності множинної впорядкованості таксономій, що дозволяє надалі реалізовувати процедури, які забезпечують формування мережевих тематичних онтологічних систем [4, 15, 23, 28, 29].

## СПИСОК ЛІТЕРАТУРИ

1. Москаленко А., Губерський Л., Іванов В. Основи масово-інформаційної діяльності. – К., 1999. – С. 271.
2. World Economic Forum. Insight Report [Електронний ресурс] / The Global Information Technology Report 2012. Living in a Hyperconnected World.– Режим доступу : [http://www3.weforum.org/docs/Global\\_IT\\_Report\\_2012.pdf](http://www3.weforum.org/docs/Global_IT_Report_2012.pdf)
3. Глушков В. М. Основы безбумажной информатики / В. М. Глушков. – М. : Наука, 1982. – 552 с.
4. Гладун В. П. Процессы формирования новых знаний [Текст] / В. П. Гладун. – София : СД «Педагог 6», 1994. – 192 с.
5. Палагин А. В. Онтологические методы и средства обработки предметных знаний / А. В. Палагин, С. Л. Кривый, Н. Г. Петренко. – [монографія] – Луганск : изд-во ВНУ им. В. Даля, 2012. – 323 с.
6. Стрижак А. Е. Операциональные характеристики онтологий / А. Е. Стрижак // Открытые информационные и компьютерные интегрированные технологии, 2014. – № 66. – С. 200–208.
7. Стрижак А. Е. Онтологические аспекты трансдисциплинарной интеграции информационных ресурсов / А. Е. Стрижак // Открытые информационные и компьютерные интегрированные технологии, 2014. № 65. – С. 211–223.
8. Стрижак О. Є. Трансдисциплінарність екологічних досліджень / О. Є. Стрижак // Екологічні науки: науково-практичний журнал. – К.: ДЕА, 2014. – № 7.– С. 25–37.

9. Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP (2-е издание) / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. – 2-е изд., перераб. и доп. – СПб. : БХВ-Петербург, 2007. – 384 с.
10. Малишевский А. В. Качественные модели в теории сложных систем. – М.: Наука. Физматлит. 1998. – 528 с.
11. Шаталкин А. И. Таксономия. Основания, принципы и правила [Текст] / А. И. Шаталкин. – М. : Товарищество научных изданий КМК, 2012. – 600 с.
12. Саати Т., Керис К. «Аналитическое планирование. Организация систем». – М.: «Радио и Связь», 1991.
13. Стрижак А. Е. Таксономические характеристики онтологических систем / А. Е. Стрижак // Бионика интеллекта, 2014. – № 2(83). – С. 24– 29.
14. Буч Г. Объектно-ориентированное проектирование с примерами применения: Пер. с англ. – М.: Конкорд, 1992. – 519 с.
15. Стрижак О. Є. Трансдисциплінарна інтеграція інформаційних ресурсів [Текст] : автореф. дис. ... д-ра техн. наук : 05.13.06 / Стрижак Олександр Євгенійович ; Нац. акад. наук України, Ін-т телекомунікацій і глобал. інформ. простору. Київ, 2014. – 47 с.
16. Клини С. К. Введение в метаматематику [Текст] / С. К. Клини. – М. : Иностранная литература, 1957. – 526 с.
17. Колмогоров А. Н., Драгалин А. Г. Математическая логика. – М.: УРСС, 2005. – 240 с.
18. Фукс Л. Частично упорядоченные алгебраические системы [Текст] / Л. Фукс. – М. : Мир, 1965. – 342 с.
19. Миркин Б. Г. Проблема группового выбора. – М.: Наука, 1974. – 256 с.
20. Саати Т. Принятие решений. Метод анализ иерархий: Пер. с англ. / Т. Саати.– М.: Радио и связь, 1989.– 316 с.
21. Gruber T. R. A translation approach to portable ontology specifications / T.R. Gruber // Knowledge Acquisition. – 1993. – Vol. 5. – P. 199–220.
22. Guarino N. Understanding, Building, and Using Ontologies // URL: <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/guarino.html>
23. Guarino N., The Ontological Level. In: Casati R., Smith N. and White G. (eds.), Philosophy and the Cognitive Sciences, Vienna: Holder-Pichler-Tempsky, 1994.
24. Мендельсон Э. Введение в математическую логику. – М. : Наука, 1971. – 320 с.
25. Коршунова С. О. Роль тезаурусного моделирования в организации терминополья «ТЕХТ-ТЕКСТ» [Электронный ресурс] / С. О. Коршунова // Вестник Иркутского государственного лингвистического университета, 2009. – № 1. – Режим доступа: <http://cyberleninka.ru/article/n/rol-tezaurusnogo-modelirovaniya-v-organizatsii-terminopolya-text-tekst>.
26. Чапайкина Н. Е. Семантический анализ текстов. Основные положения [Текст] / Н. Е. Чапайкина // Молодой ученый. – 2012. – № 5. – С. 112–115.
27. Барендрегт Х. Лямбда-исчисление. Его синтаксис и семантика: Пер. с англ. – М.: Мир, 1985. – 606 с.
28. Гаврилова Т. А. Базы знаний интеллектуальных систем [Текст] / Т. А. Гаврилова, В. Ф. Хорошевский. – СПб.: Питер, 2001. – 384 с.
29. Валькман Ю. Р. Модельно-параметрическое пространство: теория и применение : [монография] / Ю. Р. Валькман, В. И. Гриценко, А. Ю. Рыхальский. – К. : Наукова думка, 2012. – 192 с.– (Проект «Наукова книга»).– Дар видавництва «Наукова думка». НАН України.

*Стаття надійшла до редакції 22.01.17.*