

КОНТРОЛЬ, АВТОМАТИКА ТА ЕЛЕКТРОТЕХНІКА

УДК 004.942:622.286

ФОРМАЛЬНІ МЕТОДИ ІНТЕРПРЕТАЦІЇ ДАНИХ ТА ЗНАНЬ ПРО НАФТОГАЗОВІ ОБ'ЄКТИ

М.М. Демчина, В.Р. Процюк, В.І. Шекета

*ІФНТУНГ, 76019, м. Івано-Франківськ, вул. Карпатська, 15, тел. (03422) 42127,
e-mail: demचना@i.ua*

Виконано аналіз формально-логічних підходів до представлення та інтерпретації даних і знань про нафтогазові об'єкти у вигляді дослідження функціональності фреймових, продукційних та семантичних представлень. Представлено структуру продукційного правила, умова якого уможливує використання сполучників І та АБО, а логічна поведінка правила описується коефіцієнтами впевненості. Визначено умову формування продукційних правил на основі об'єктів, компараторів, доменів, атрибутів, та опційних значень. Показано, що введення ієрархії доменів для нафтогазової предметної області дозволяє виконати структурування та концептуалізацію логічного висновку відповідно до введеного впорядкування множини цілей. Введено структуру числових, логічних та категорійних атрибутів, що забезпечує генерування послідовності діалогових питань системою. Специфіковано характеристики доменної організації предметної області, що визначається особливостями логічного висновку, способом специфікації та впорядкування правил. Побудовано семантичні представлення для об'єктів предметної області, шляхом визначення відповідних семантичних зв'язків між об'єктами та атрибутами, об'єктами та правилами, атрибутами та значеннями, правилами та доменами. Обґрунтовано доцільність імплементації гібридного підходу для нафтогазової предметної області у вигляді виділених структурних елементів та їх композицій.

Ключові слова: нафтогазовий об'єкт, атрибут, знання, інтелектуальні системи, бази знань, інтерпретація даних та знань, фрейми, семантичні мережі, продукції.

Выполнен анализ формально-логических подходов представления и интерпретации данных и знаний о нефтегазовых объектах путём исследования функциональности фреймовых, продукционных и семантических представлений. Представлена структура продукционного правила, условие которого позволяет использовать логические операторы И, ИЛИ, а логика поведения правила описывается коэффициентами уверенности. Установлено условие формирования продукционных правил на основании объектов, компараторов, доменов, атрибутов и опционных значений. Показано, что введение иерархии доменов для нефтегазовой предметной области позволяет совершать структурувание и концептуализацию логического вывода согласно использованного способа упорядочивания множества целей. Введена структура числовых, логических атрибутов, а также атрибута категории, что обеспечивает генерирование последовательности диалоговых вопросов системой. Специфицированы характеристики доменной организации предметной области, что определяется особенностями логического вывода, способом спецификации и упорядочивания правил. Построены семантические представления для объектов предметной области путём определения соответствующих семантических связей между объектами и атрибутами, объектами и правилами, атрибутами и значениями, правилами и доменами. Обосновано целесообразность имплементации гибридного подхода для нефтегазовой предметной области в виде выделенных структурных элементов и их композиций.

Ключевые слова: нефтегазовый объект, знания, интеллектуальные системы, базы знаний, интерпретация данных и знаний, фреймы, семантические сети, продукция.

The analysis of formal-logical approaches to representation and interpretation of data and knowledge about oil-and-gas objects in the form of functionality exploration for frame-based, semantic-based and production-based representation forms is done. The production rule structure is presented in form that allows using AND, OR – operators, and with certainty factors as descriptions means for rule behavior modeling. The condition for production rules forming routine based on objects, comparators, domains, attributes and optional values is defined. It is shown, that introducing domain hierarchies for oil-and-gas related domain allows structuring and conceptualization of logical inferences accordingly to introduced goals ordering set. The structure of numerical, logical and categorical attributes is proposed, that allows the system to generate the sequences of questions in dialog form. The domains implementation is featured for subject area based on peculiarities of logical inferences, ordering rules and specification rules. The semantic presentations for related domain objects is constructed, by the means of defining corresponding semantic relations between objects and attributes, objects and rules, attributes and values, rules and

domains. The implementation expediency for hybrid approach for oil-and-gas related-domain in the form of selected structural elements and theirs compositions is grounded.

Keywords: oil-and-gas object, attribute, knowledge, intellectual systems, knowledge bases, data interpretation and knowledge, data frames, semantic networks, production rules.

В інтелектуальних системах використовуються ряд методів представлення знань: логічний, графовий, метод вектор-функцій [1-5]. Логічний метод базується на використанні числення предикатів першого порядку для формування моделі предметної області, методів простору станів, методів доведення теорем та евристик як механізмів пошуку рішень.

В основі семантичних моделей лежать ідеї семантичних мереж, що складаються з множини об'єктів які представляють факти, об'єкти, властивості і зв'язки між ними (відношень рівності, належності, причинно-наслідкові зв'язки і т. д.). За своєю сутністю семантична мережа є орієнтованим графом. Залежно від типів зв'язків, моделі поділяються на класифікуючі, функціональні і сценарії. На основі цих зв'язків можливо представляти ієрархічні структури, обчислювальні процедури та інші складні залежності.

Продукційна модель містить елементи логічних та семантичних представлень. Процес нагромадження знань у таких моделях відбувається за рахунок трансформації семантичної мережі в результаті застосування правил висновку – продукцій. В результаті цього змінюються фрагменти мережі і видаляються непотрібні частини. Дана форма також відрізняється тим, що процедурні знання чітко відокремлені від декларативних і описуються в іншим спосіб.

Фрейм характеризується тим, що має жорстку структуру елементів знань. Відповідно таке представлення подібне до представлення класів об'єктно-орієнтованого підходу. Окрім того кожен фрейм є проіменованим контейнером із набором складових частин – теж проіменованих елементів певного виду. Тут складовими можуть бути: числа, текст, властивості, умови, процедури та інші фрейми. Зв'язки між фреймами описуються фреймом спеціального виду.

Таким чином, кожній формальній моделі властиві власний спосіб опису знань, однак всі вони не позбавлені недоліків: недостатня універсальність, складність отримання нових знань, можливість отримання суперечливих знань, складність розширення моделі, надмірний розмір моделі, відсутність наочності в представленні знань та ряд інших. Для нафтогазової предметної області **невирішеною** є задача комплексної інтелектуальної обробки даних та знань про нафтогазові об'єкти на основі використання існуючих моделей знань.

Метою даної статті є розроблення формально обґрунтованого методу інтерпретації даних та знань про нафтогазові об'єкти, що виділяється на різних етапах життєвого циклу родовища.

Означення 1. Знаннями про нафтогазовий об'єкт будемо вважати інтерпретацію в термінах штучного інтелекту сукупності даних, понять, фактів та закономірностей нафтогазової

предметної області, що об'єднуються певною ознакою яка уможливіє представлення цілісної структури. В такому розумінні предметні знання можна розглядати як моделі інтелектуальної поведінки експертів предметної області (геологів, фахівців з проблем видобутку та транспортування нафти і газу тощо).

Початкову формалізацію знань предметної області (*SDK – Subject Domain Knowledge's*) представимо у вигляді залежності множини задач предметної області (*SDT – Subject Domain Tasks*) і доступних ресурсів предметної області (*SDR – Subject Domain Resources*) у вигляді множини правил (*Rules – Subject Domain Rules*), а саме:

$$Knowledges^{set} = (SDT^{set}, SDR^{set}, Rules^{set}). \quad (1)$$

Проте, незалежно від способу представлення та функціональності наявність входження знань в певній інформаційній системі не можуть характеризувати її як систему інтелектуального рівня. Вирішення задачі інтелектуалізації в даному випадку полягає у створенні баз знань як якісної систематизації та класифікації зв'язків та відношень між об'єктами предметної області.

Означення 2. Базою знань нафтогазової предметної області (*KB – Knowledge Base*) будемо називати структуру представлення та управління наборами фактів, правил та процедур логічного висновку на множині нафтогазових об'єктів. Класичне рішення задачі розроблення баз знань вимагає: розроблення програмних структур та засобів для представлення та зберігання знань; збір та розміщення знань предметної області в базі.

За допомогою чіткої математичної основи категорій та функторів можна досліджувати семантичний зміст правильно з математичної точки зору (шляхом побудови різноманітних моделей представлення знань), що необхідно для формалізації знань, побудови баз знань та інтелектуальної підтримки прийняття рішень.

Логічна модель призначена для розв'язання простих однорідних задач. Вона базується на алгебрі тверджень та предикатів, системах, аксіомах та методах висновку. Відповідно у кожній логіці наявний синтаксис, що визначає правила побудови синтаксично та семантично коректних висловлювань, і семантики як правил інтерпретації логічних тверджень.

В основі логічного представлення знань (*LKP – Logical Knowledge's Presentation*) лежить формальна теорія, що виражається кортежем:

$$LKP = \left(BS^{set}, \left(F : bs \in BS^{set} \rightarrow \right)^{set}, TF, FR \right), \quad (2)$$

де: BS^{set} – скінченна множина базових символів; F^{set} – множина формул у вигляді синтаксичних правил, за допомогою яких з елементів BS^{set} формуються коректні твердження; TF – множина істинних формул; FR – множина відношень між формулами (правилами).

Крім того, для множини BS^{set} повинні існувати процедури верифікації щодо належності елементів та входжень синтаксично правильних сутностей.

Основні проблеми імплементації логічної моделі полягають в усуненні суперечностей даних та наявності процедур (правил) оновлення входжень кортежу логічного представлення.

Логіка тверджень розглядає судження як єдине ціле, що має лише властивість істинності. У разі необхідності розробити логічну модель знань зручно використовувати логіку предикатів. Логіка предикатів більше підходить для відтворення логічних міркувань, оскільки дозволяє врахувати будову і зміст простих висловлювань.

Означення 3. Численням предикатів вважають формальну мову подання логічних та інших видів відношень, результатом яких є булева величина.

Центральним твердженням у численні предикатів є те, що кожне просте висловлювання містить в собі інформацію про деякий об'єкт і його властивості. Числення предикатів має потужний і зрозумілий механізм рекурсивного висновку. Суттєвою якістю є також широкий набір виконуваних операцій, зокрема квантор існування (\exists) і квантор узагальнення (\forall), що дає змогу описувати загальні властивості множини об'єктів.

Відповідно базу знань на основі предикатів можна представити у вигляді:

$$KB = (A, SR^{set}, TF^{set}, IR^{set}), \quad (3)$$

де: A – алфавіт (сукупність використовуваних символів); SR^{set} – множина синтаксичних правил; TF^{set} – початкова множина істинних формул; IR^{set} – множина правил висновку для побудови похідних формул.

Інтерпретація логіки предикатів наступна. Нехай $Object^{set}$ – деяка множина об'єктів і O_1, O_2, O_3 – елементи цієї множини, відповідно логічні висловлювання про ці об'єкти позначатимемо як $State(O_1), State(O_2), State(O_3), State(O_4)$. Згідно з початковим припущенням дані висловлювання будуть істинними або хибними.

Нехай x – випадковий елемент з $Object^{set}$, тобто $x \in Object^{set}$, тоді $State(x)$ позначатиме висловлювання, що стає визначеним коли x замінюється конкретним елементом з $Object^{set}$. З іншого боку $State(x)$ представлятиме булеву функцію на множині $Object^{set}$. Відповідно не-

визначені функції будемо розглядати як предикатні, при чому в таких предикатах змінні можуть бути як з одного алфавіту, так і з різних:

$$[Object^{set}]^n \rightarrow Bool^{set}, \quad (4)$$

$$[Object^{set}]^1 \times [Object^{set}]^2 \times \dots \times [Object^{set}]^n \rightarrow Bool^{set}, \quad (5)$$

де: $[Object^{set}]^1, [Object^{set}]^2, \dots, [Object^{set}]^n$ – алфавіти; $Bool^{set}$ – булева множина.

Застосування семантичних моделей в інженерії знань бере свій початок від методів семантичного аналізу природньомовних текстів, що з розвитком мов програмування призвело до створення їх семантик, а на сьогоднішній день концепція семантизації WEB розглядається як найбільш актуальний тренд розвитку WEB та Internet технологій.

Означення 4. Семантичною мережею (*Semantic Network*) вважають спосіб представлення інформації, яка базується на графовій інтерпретації. Вважається, що предметну область можна представити у вигляді семантичної мережі якщо її модель можна звести до набору двомісних предикатів, які згідно свого означення дозволяють подання у вигляді графів, де ребрам відповідають предикати, а вершинам аргументи. Таким чином семантичну мережу можна розглядати як орієнтований граф з таких складових:

1) множина вершин (*Vertices Set*) яка використовується для опису та представлення даних, об'єктів.

2) множина ребер (*Edges Set*) дозволяє моделювати відношення між об'єктами.

Таким чином знання що базуються на семантичній мережі можна формально представити кортежем:

$$K^S = (VS, A^{VS}, AVS, RTS, PS[VS]), \quad (6)$$

де: K^S (*Knowledges^{set}*) – множина знань;

VS (*VerticesSet*) – множина вершин графа;

A^{VS} (*Attributes^{VerticesSet}*) – множина атрибутів приписаних вершинам;

$A.VS$ (*Attributes.ValuesSet*) – множини значень атрибутів;

RTS (*RelationsTypesSet*) – множина типів зв'язків, введених в мережі;

$PS[VS]$ (*Projections.Set[VerticesSet]*) – множина проєкцій для вершин по видам зв'язків.

Відповідна структуризація зв'язків задаватиметься кортежем:

$$RTS = (LR, SemR, HR, FR, SetR, LR, AR, CR, QR), \quad (7)$$

де: LR (*LogicalRelations*) – логічні зв'язки,

$LogicalRelations = \{\wedge, \vee, \Rightarrow, \Leftrightarrow\}$;

$SetR$ (*SetRelations*) – множинні зв'язки,

$SetRelations = \{SubSets, ClassElements, ClassExamples\}$;

HR (HirerarchicalRelations) – ієрархічні зв'язки,

$HirerarchicalRelations = \{Inherrtance, SubOrdering\}$;

SemR (SemanticRelations) – семантичні зв'язки,

$SemanticRelations = \{VerticesGoals, GoalsOrdering\}$;

FR (FunctionalRelations) – функціональні зв'язки,

$FunctionalRelations = \{ObjectProperties\}$;

LR (LinguisticRelations) – лінгвістичні зв'язки,

$LinguisticRelations = \{LinguisticValues\}$;

AR (AttributiveRelations) – атрибутивні зв'язки, $AttributiveRelations = \{AttributeValues\}$;

CR (CasualRelations) – казуальні зв'язки,

$CasualRelations = \{VerticesEvents, ImplicationsSet\}$;

QR (QuantificationRelations) – квантифікаційні зв'язки,

$QuantificationRelations = QuantorSet = \{\forall, \exists\}$.

Відповідно завдяки однорідній структурі семантичні мережі дозволяють спростити процес представлення, оновлення та підтримки знань, а також організацію доступу до них та імплементацію процедури висновку. Також, класично розрізняють причинно-наслідкові, асоціативні, екстенсивні та інтенсивні мережі. Відповідно, якщо семантичні мережі мають тільки один тип відношень то вони розглядаються як однорідні, інакше – як неоднорідні. Найбільш простими для відображення на графі природно є бінарні відношення. Відповідно при зростанні розмірності відношень, зростає також складність їх ефективного графового представлення. Поширеними типами відношень є відношення наслідування, причинної наслідковості, поширення властивості, поширення значень, схожості та елементної належності.

В загальному випадку вважається, що моделювання довільного відношення найбільш доцільно виконати у вигляді концептуального графу, який формально можна описати як

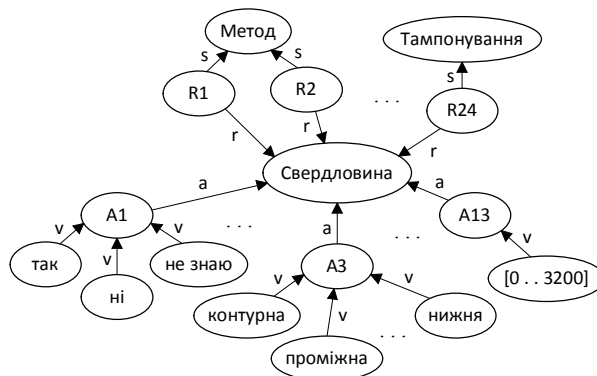
$$CG = \left\langle \begin{matrix} Concepts^{set}, Relations^{set}, \\ Markers^{set} [Constraints^{set}, Conditions^{set}] \end{matrix} \right\rangle, (8)$$

де: $Concepts^{set}$ – множина концептів предметної області; $Relations^{set}$ – множина відношень введених в предметній області; $Markers^{set}$ – множина маркерів з введеними підмножинами обмежень ($Constraints^{set}$) та умовами ($Conditions^{set}$).

Засобами семантичної мережі є можливим виконувати структурування множини відношень, виділяючи деяку підмножину первинних відношень, що розміщуються у вершинах і похідних відношень, які розміщуються у відповідних вершинах, зв'язаних з первинними.

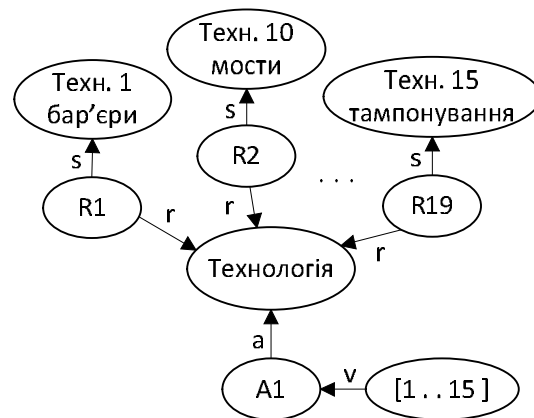
Для представлення знань за допомогою семантичної мережі введемо поняття нафтога-

зових об'єктів, атрибутів, значень, правил та доменів. У спроектованій семантичній мережі використовуватимемо наступні позначення: *a* – семантичні зв'язки між нафтогазовим об'єктом та атрибутами; *r* – семантичні зв'язки між нафтогазовим об'єктом та правилами; *v* – семантичні зв'язки між атрибутами та значеннями; *s* – семантичні зв'язки між правилами та доменами.



$A1 = \langle \text{«Є обводненість»} \rangle$, $A2 = \langle \text{«Є залишкова нафтонасиченість»} \rangle$, ..., $A13 = \langle \text{«Пластовий тиск, \%»} \rangle$ – множина атрибутів нафтогазового об'єкта; $R1 = \langle \text{«Метод бар'єри»} \rangle$, $R2 = \langle \text{«Метод екрани»} \rangle$, ..., $R24 = \langle \text{«Технологія 9. Тампонування»} \rangle$ – множина правил для нафтогазового об'єкта

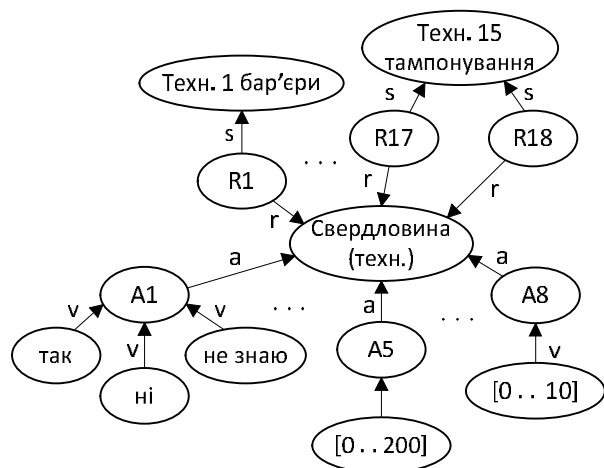
Рисунок 1 – Семантичне представлення першого виду для об'єкта «Свердловина»



$A1 = \langle \text{«Номер»} \rangle$ – атрибут нафтогазового об'єкта; $R1 = \langle \text{«Технологія 1. Бар'єри (техн.)»} \rangle$, $R2 = \langle \text{«Технологія 10. Мости (техн.)»} \rangle$, ..., $R19 = \langle \text{«Технологія 15. Тампонування (техн.)»} \rangle$ – множина правил

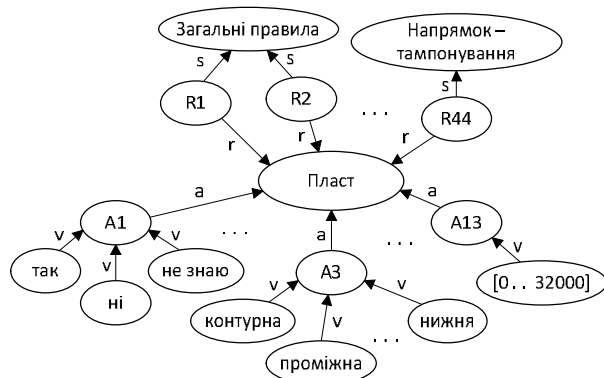
Рисунок 2 – Семантичне представлення першого виду для об'єкта «Технологія вилучення»

Як було показано в роботах [7-9], найбільш поширеним способом представлення знань для експертних систем навчального та комерційного застосування є метод продукційних правил. Відповідно досліджується цілий клас систем на основі правил (*Rules Based System*). Популярність



A1=«Герметичність», A2=«Пакування», ..., A8=«Перфорація» – множина атрибутів; R1=«Технологія 1. Бар'єри (техн.)», R2=«Технологія 10. Мости (техн.)», ..., R18=«Технологія 15. Тампонування (техн.)» – множина правил

Рисунок 3 – Семантичне представлення другого виду для об'єкта «Свердловина»



A1=«Є обводненість», A2=«Є залишок нафти», ..., A13=«Пластовий тиск, %» – множина атрибутів; R1=«Загальні правила 1», R2=«Загальні правила 2», ..., R44=«Технологія 9. Тампонування» – множина правил

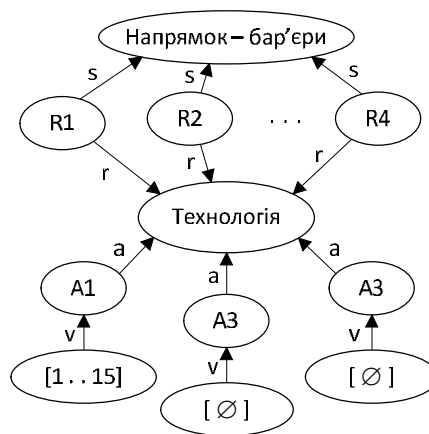
Рисунок 4 – Семантичне представлення для об'єкта «Пласт»

даного підходу зумовлюється його вираженою логічною орієнтацією в термінах класичних пропозиційних та розширених логік і природністю середовища організації логічного висновку. В даному підході початкова формалізація предметної області виконується у вигляді набору фактів та правил:

$$SubjectDomainModel = Facts^{set} \cup Rules^{set} \quad (9)$$

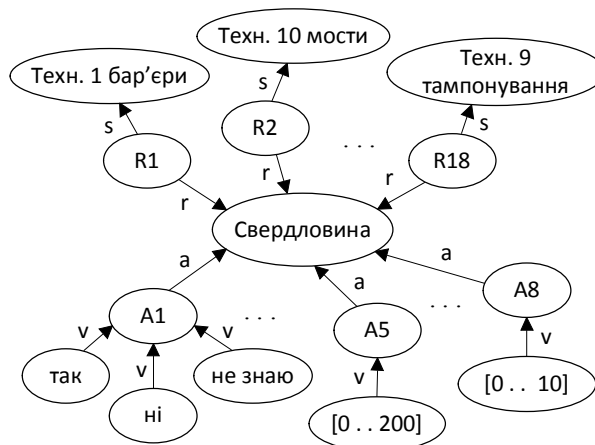
Означення 5. Фактом предметної області вважається вираження істинності або хибності певного висловлювання. Традиційно факти описують властивості:

$$Facts^{set} = \{Object_i.Property_j\}_{i,j \in N} \quad (10)$$



A1=«Номер», A2=«Кількість невідомо», A3=«Кількість неправильно» – множина атрибутів; R1=«Бар'єри 1», R2=«Бар'єри 2», ..., R4=«Технологія 15. Тампонування» – множина правил

Рисунок 5 – Семантичне представлення другого виду для об'єкта «Технологія вилучення»



A1=«Герметичність», A2=«Пакування», ..., A8=«Перфорація» – множина атрибутів; R1=«Технологія 1. Бар'єри», R2=«Технологія 10. Мости (техн.)», ..., R18=«Технологія 9. Тампонування (техн.)» – множина правил.

Рисунок 6 – Семантичне представлення третього виду для об'єкта «Свердловина»

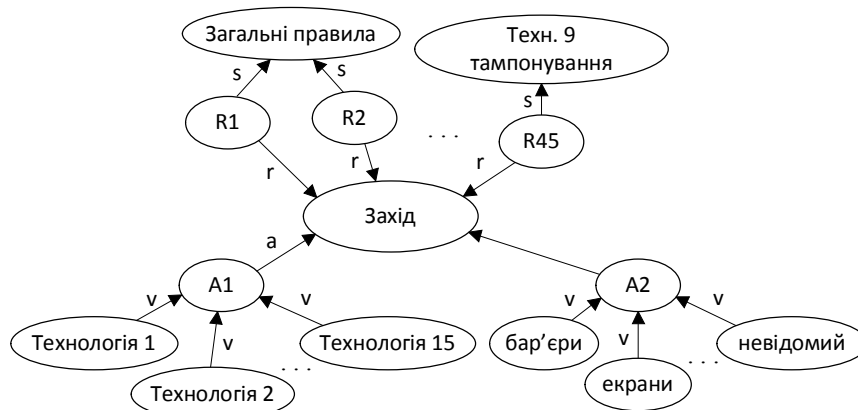
Означення 6. Правилком вважають вираження зв'язків залежностей та слідувань на множині висловлювань та фактів:

$$RulesSet = \left\{ \left[Facts^{set}, Statements^{set} \right], \left[Dependencies, Implications \right] \right\} \quad (11)$$

$$Rules^{set} \rightarrow \{ExpertAdvices, Inferences\} \quad (12)$$

Відповідно множина правил дозволяє системі будувати експертні поради та стратегії висновку:

Функціональність системи на основі правил полягає у верифікації висловлювань засобами множини правил. Відповідно висловлю-



$A1$ = «Висновок», $A2$ = «Напрямок» – множина атрибутів; $R1$ = « Загальні правила 2», $R2$ = « Загальні правила 3», ..., $R45$ = «Технологія 9. Тампонування (техн.)» – множина правил

Рисунок 7 – Семантичне представлення для об'єкта «Захід»

вання верифіковані як істинні додаються до множини фактів:

$$\{Statements^{set}\} \xrightarrow{Rules^{set}} Facts^{set}. \quad (13)$$

В якості переваг підходу на основі правил слід виділити: адаптивність, ієрархічність, інтуїтивність, природність, шаблонність, модульність. В якості недоліків виділяють: складність ієрархічного впорядкування множини правил; складність процедури логічного висновку при великій кількості правил. На формальному рівні продукційну модель можна зобразити у вигляді:

$$PM = \left\langle \begin{matrix} ClassDescription, ActivationCondition, \\ H \rightarrow B, PostCondition \end{matrix} \right\rangle. \quad (14)$$

На множині продукцій кожне входження виділяється унікальним іменем з власною областю застосування яка позначається як домен. Таким чином, вся база знань структурується, що дає змогу оптимізувати доступ до бази даних і відповідно задати необхідні обмеження:

$$KB = \{Scopes_i\}_{i=1..n, n \in N}. \quad (15)$$

Відповідно умова застосування ядра продукцій представляється в предикатній формі з додатково можливою реалізацією постумов. Структура ядра продукції може бути імплікаційною або умовною, що дозволяє визначати стратегії детермінованої або недетермінованої поведінки. Стратегія детермінованості визначає обов'язковість виконання дії, а стратегія недетермінованості дозволяє ймовірнісну та можливісну інтерпретацію.

Продукційні правила представляються у вигляді:

$$IF \langle Conditions^{set} \rangle THEN \langle Actions^{set} \rangle. \quad (16)$$

Процедура оцінювання умов здійснюється на введеній множині фактів:

$$Evaluation(Condition^{set}, Facts^{set}). \quad (17)$$

Відповідно кількість виконуваних дій залежить від кількості задоволених фактів предметної області. Побудова умови виконується як

процедура інтерпретації твердження на множині фактів в процесі здійснення логічного висновку з метою задоволення визначеної цілі:

$$C \rightarrow IR(S^{set}) \rightarrow \{SR(F^{set}), A^{set}(I(D, GS))\}, \quad (18)$$

де: C (Condition) – умова;

IR (Interpretation Runtines) – процедури інтерпретації; S^{set} ($Statements^{set}$) – множина тверджень; SR (Search Runtines) – пошукові процедури; F^{set} ($Facts^{set}$) – множина фактів; A^{set} ($Actions^{set}$) – множина дій; I (Inferences) – процедури логічного висновку; D (Deducing) – процедури дедукції;

GS (Goal Satisfaction) – процедури задоволення цілі.

Відповідно оптимізація продукційної моделі може бути здійснена шляхом введення деякого способу впорядкування множини продукційних правил або шляхом додавання множини пріоритетів до правил продукції.

$$ProductionModel(Ordering, Priority_{PM}^{set}). \quad (19)$$

Таким чином, множини продукції можуть описувати певні стани предметної області, представлені у формі секцій, або описувати комплексні рішення як процес задоволення цілі.

Фреймовий підхід до представлення знань базується на теорії фреймів. Дана модель часто розглядається як добре структурована, об'єктно-орієнтована модель предметної області. В найпростішій інтерпретації фрейми – це структури для зберігання даних.

Означення 7. Фреймом вважають структурований фрагмент знань, що описує певну ситуацію предметної області. Суть класичних результатів в даній області зводиться до того, що фреймова інформація категоризується по кількох рівнях: перший рівень описує спосіб використання фрейму; другий рівень – очікувана поведінка фрейму; третій рівень – нестандартні ситуації і т.д. Таким чином, фрейм можна розглядати як контейнер для знань і в той же час як входження бази даних. Крім того, фрейм представляє вузькоспеціалізовані зв'язані знан-

ня по виділеній темі з частиною входжень визначених як значення по замовчуванню. Важливим елементом також є можливість застосування ієрархії фреймів, в якій введена операція наслідування як і об'єктно-орієнтованих реалізацій в імперативних мовах програмування.

На формальному рівні фреймову модель доцільно представляти кортежем виду:

$$FM = \langle FS, \{F_i.SS\}, DT, VR, RS, IR, \rangle, (20)$$

де: FM (*FrameModel*) – фреймова модель;
 FS (*FrameSet*) – множина фреймів;
 $F_i.SS$ (*Frame_i.SlotSet*) – множина слотів i -того фрейму; DT (*DataTypes*) – типи даних; VR (*ValuesRange*) – діапазон значень; RS (*RoutinesSet*) – множина процедур; IR (*InheritanceRelations*) – відношення наслідування.

Відповідно до ситуації предметної області система вибирає певний фрейм з бази, що відповідає даному класу ситуації, виконує його узгодження з описом ситуації шляхом актуалізації певних даних з відповідних наборів, що призводить до отримання максимально адаптивного представлення фрейму. При такому підході система використовує два типи фреймів: прототипні фрейми із заданими значеннями комірок за замовчуванням та ініціалізовані фрейми з ініціалізованими значеннями комірок відповідно до фактичного стану предметної області. Такий опис дозволяє отримати деталізацію формальної структури фрейму у вигляді:

$$FS = \left\{ \left\langle Slot_1, Value_1 \right\rangle, \left\langle Slot_2, Value_2 \right\rangle, \dots, \left\langle Slot_k, Value_k \right\rangle \right\}, (21)$$

де: $Slot_k$ – k -й слот;
 $Value_k$ – значення k -го слоту.

Деталізація значень слотів може бути зведена до:

$$Slot_i.Value_i = \left\{ \left\langle Slot_i.Attr_{j_1}, Attr_{j_1}.Values^{set} \right\rangle, \dots, \left\langle Slot_i.Attr_{j_n}, Attr_{j_n}.Values^{set} \right\rangle; Ref_1, \dots, Ref_i \right\}, (22)$$

де: $Slot_i.Attr_j$ – j -й атрибут i -го слоту;
 $Attr_j.Values^{set}$ – множина значень j -го атрибуту;
 Ref_i – посилання на інші слоти або фрейми.

Структуру слоту можна визначити формальною структурою виду:

$$SS = \left\{ SlotName, DateType, DefaultValue, InheritanceFeature, Routines, Rules, Pictures, Remarks, Assumes \right\}, (23)$$

де: $SlotName$ – ім'я слоту; $DateType$ – тип даних; $DefaultValue$ – значення за замовчуванням; $InheritanceFeature$ – особливість наслідування; $Routines$ – процедури; $Rules$ – правила; $Pictures$ – зображення; $Remarks$ – коментарі; $Assumes$ – припущення щодо стратегії поведінки.

Означення 8. Фреймовою процедурою називають процедуру, що запускається автоматично при досягненні певної умови (в основному при операціях з вмістимим слотів).

Розрізняють процедури, що запускаються з поточного фрейму і сервісні процедури, які запускаються з інших фреймів.

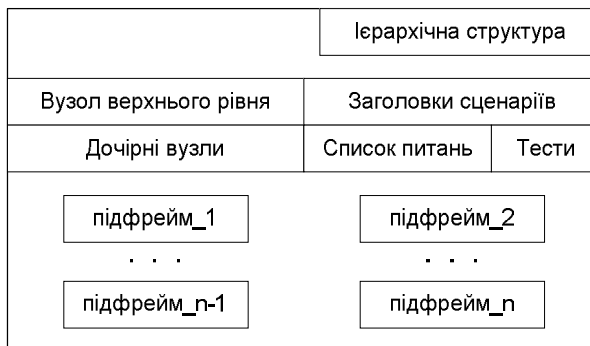


Рисунок 8 – Ієрархічна фреймова структура

Означення 9. Фреймовою системою (*Frame System*) називають сукупність фреймів з ієрархічною структурою, яка утворює одну з можливих моделей предметної області.

Для представлення знань у нафтогазовій предметній області можна скористатися системою фреймів, яка складається з таких компонентів: об'єкт, категорійний атрибут, логічний атрибут, числовий атрибут, домен, цілі, правило.

Об'єкт складається з імені та включає атрибути, які можуть бути одного із трьох типів: категорійного, логічного і числового.

Як значення категорійного атрибуту використовується послідовність елементів і його загальний вигляд можна зобразити у наступній символічній формі:

$$EnumeratedAttr = \left\{ Value_1, Value_2, \dots, Value_i \right\}_{i=1..n, n \in N} \quad (24)$$

Логічний атрибут, у свою чергу, може приймати одне з трьох можливих значень: *True*, *False* та *Unknown* і у символічній формі може бути представленим як:

$$BooleanAttr = \{ True, False, Unknown \} \quad (25)$$

За допомогою числового атрибуту є можливість зберігання числових значень з деякого діапазону, який задається у наступній формі: початкове значення, кінцеве значення та значення за замовчуванням.

$$NumericalAttr = \left[\begin{matrix} [Value_{min} .. Value_{max}] \\ DefaultValue \end{matrix} \right] \quad (26)$$

Усі з перелічених атрибутів повинні мати унікальне ім'я, короткий опис, та запитання яке буде відображатися системою під час консультації.

Для зв'язування об'єктів у смислові конструкції використовуються правила, на основі яких системою формується логічний висновок. Кожне правило повинне мати унікальне ім'я. За допомогою коефіцієнта впевненості можна задати спосіб застосування правила при істинній

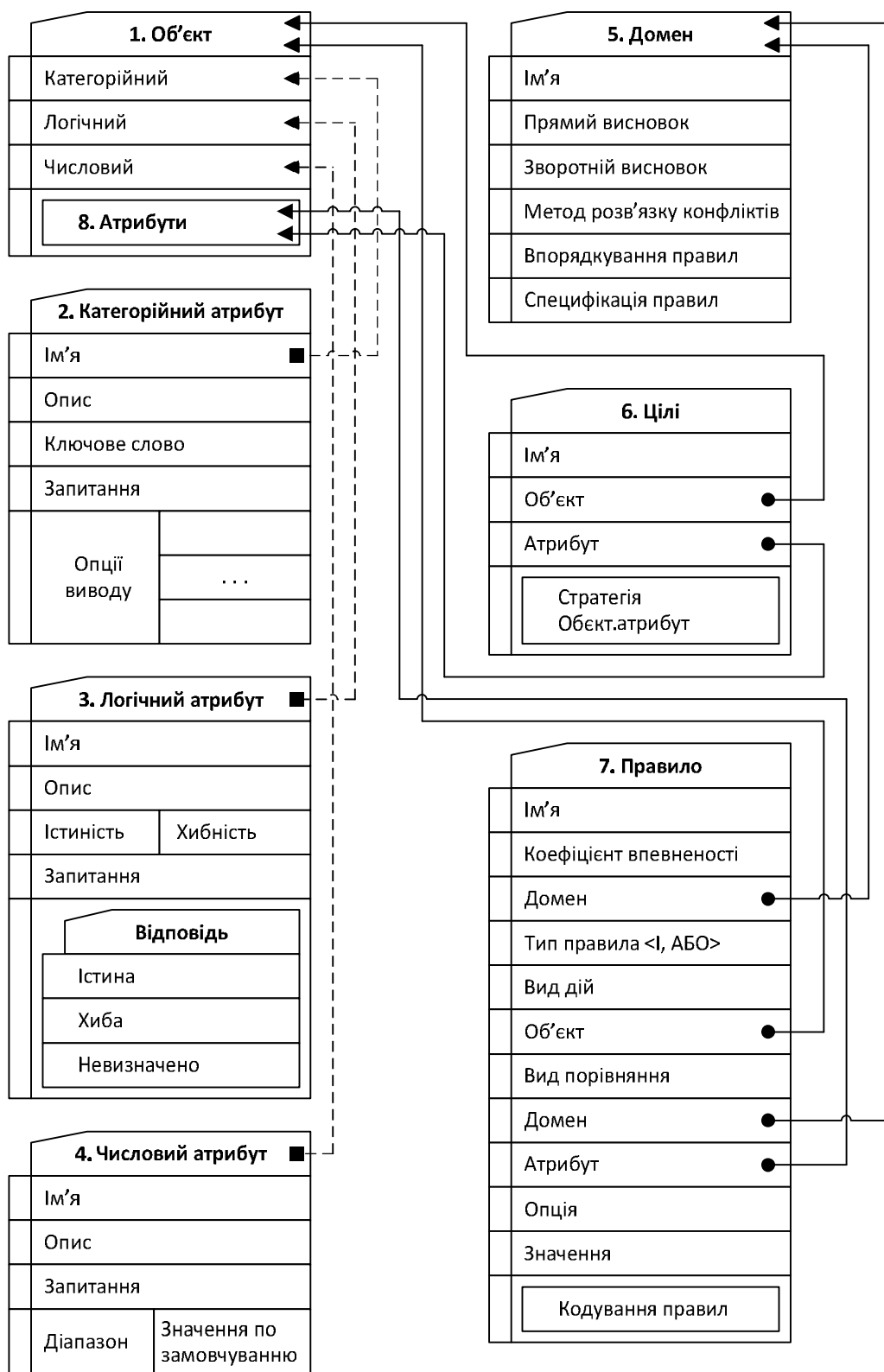


Рисунок 9 – Представлення знань за допомогою зв'язаної системи фреймів

передумові. Об'єкти можуть зв'язуватися у смислові конструкції за допомогою двох типів логічних зв'язків І та АБО, відповідно бінарна операція І використовується для отримання логічного добутку між двома об'єктами, а операція АБО – застосовується для логічного сумування об'єктів.

Умова формується за допомогою об'єкта, компаратора, домену, атрибуту, опції або значення у випадку числової властивості. Умова може складатися з довільної кількості передумов. В результаті виконання однієї з двох логічних операцій, системою виконується певний вид дії, який є логічним завершенням умови.

Прикладом правила з нафтогазової предметної області може бути правило виду:

Якщо Розмір тріщин > 10 мм
I Пластова температура <= 40 °C
I Мінералізація води > 10 г/л
Тоді активувати Технологія 11.
 Бар'єри (100)

Правила активуються у домені, який також повинен бути ідентифікованим унікальним іменем. За допомогою опцій виводу є можливим вказати, який тип висновку буде використовуватись системою під час консультації, а блок вирішення конфліктів дозволяє визначити метод розв'язку конфліктів, спосіб впорядкування правил та специфікацію правил.

Таким чином, ієрархія фреймів дає змогу виконувати таксономізацію та онтологізацію предметної області шляхом розміщення на верхніх рівнях більш значимих понять, які деталізуються на нижчих рівнях. Така ієрархія може функціонувати відповідно в статичному та динамічному режимі.

Висновки. Отже в даній статті було введено ключові поняття представлення знань, виділено відмінні ознаки знань та даних нафтогазової предметної області, розглянуто класифікацію знань за змістом та типом і проаналізовано існуючі підходи до отримання знань в предметній області. Досліджено методи, технології та підходи для отримання та подання знань в інформаційних інтелектуальних системах, описані основні моделі для реалізації знань у вигляді цілісних структур. Визначено поняття та властивості знань предметної області, їх типи, особливості та загальну функціональність інформаційних інтелектуальних систем на основі баз даних та знань нафтогазової предметної області. В результаті виділено основні вимоги до шуканої моделі представлення знань, опис та функціональність її очікуваних складових на рівні прототипів. Таким чином, реалізація ефективного алгоритму роботи з моделлю знань може бути актуалізована виділеними концептами нафтогазової предметної області. В якості базових складових реалізації гібридного підходу представлення знань про нафтогазові об'єкти виділено логічну, семантичну, продукційну та фреймову модель.

Подальші дослідження даного напрямку будуть спрямовані на практичну імплементацію до виділених гібридних підходів представлення знань в інформаційній інтелектуальній системі інтерпретації нафтогазових об'єктів.

Література

- 1 Поспелов Д.А. Данные и знания. Искусственный интеллект / Д.А. Поспелов. – М.: Радио и связь, 1990. – 300 с.
- 2 Гуц А.К. Математическая логика и теория алгоритмов: учеб. пособ. / А.К. Гуц. – Омск: Наследие, 2003. – 108 с.
- 3 S. Jajodia, P. Samarati, V. Subrahmanian. A Logical Language for Expressing Authorizations. IEEE Computer Society Press, Los Alamitos, 1997. – P. 31-42.
- 4 P. Harmon. The Size of the Commercial AI Market in the US. Intelligent Software Strategies. - v.10. – n.1. – 1994. – P. 1-16.
- 5 Лорьер Ж.Л. Системы искусственного интеллекта / Ж.Л. Лорьер. – М.: Мир, 1991. – 400 с.
- 6 Vitaliy Melnyk, Roman Vovk, Mykola Demchyna. Frame Based Approach to Construction of Intelligent System for Student Knowledge Control // Сучасні проблеми радіоелектроніки, телекомунікацій та комп'ютерної інженерії: матеріали 10-тої Ювілейної міжн. наук.-тех. конф., 23-27 лютого 2010 р. – Львів, 2010. – С. 248-250.
- 7 Мельник В.Д. Прийняття рішень при модифікації запитів в системі дистанційного навчання / В.Д. Мельник, В.І. Шекета, М.М. Демчина // Теоретичні та прикладні аспекти побудови програмних систем: VI Міжн. конф., 9-10 грудня 2009 р. – К., 2009 – С. 149-156.
- 8 Демчина М.М. Застосування концептуальних елементів експертних систем для контролю знань / [Демчина М.М., Шекета В.І., Федорак Р.М.] // Комп'ютерні науки та інформаційні технології: матеріали IV Міжн. наук.-тех. конф. CSIT-2009., 15-17 жовтня 2009 р. – Львів, 2009 – С. 153-156.
- 9 Юрчишин В.М. Інформаційне моделювання нафтогазових об'єктів: монографія / Юрчишин В.М., Шекета В.І., Юрчишин О.В. – Івано-Франківськ: Вид-во Івано-Франківського нац. техн. ун-ту нафти і газу, 2010. – 196 с.

*Стаття надійшла до редакційної колегії
 17.03.11*

*Рекомендована до друку професором
 Д.Ф. Тимківим*