

Інформаційні технології

УДК 004.82:622.286

DOI: 10.31471/1993-9973-2018-3(68)-46-59

ЙМОВІРНІСНА ОЦІНКА РЕЗУЛЬТАТІВ ІНТЕРПРЕТАЦІЇ ДАНИХ ТА ПАРАМЕТРІВ ГЕОФІЗИЧНИХ ДОСЛІДЖЕНЬ

Л.М. Гобир, Р.Б. Вовк, Л.О. Потеряло, В.І. Шекета

Івано-Франківський національний технічний університет нафти і газу;
вул. Карпатська, 15, м. Івано-Франківськ, 76019, e-mail: vasyshketa@gmail.com

Розроблено кейс-базовану методику набуття знань у ході імплементації діалогу «експерт - інтелектуальна система». Запропоновано стратегію темпорального представлення видобутих кейсів даних та знань про предметну область. Розроблено методику ініціалізації функції належності методом формальних оцінок кейсів, яка використовується для обробки результатів діалогу з групою фахівців з контенту. Застосовано розроблений формально-логічний апарат до оброблення кейсів реальних даних предметної області. Проаналізовано і узагальнено процес прогнозування вуглеводневих запасів інтелектуальною системою. Запропоновано практичну реалізацію архітектури бази кейсів інтелектуальної системи для прогнозування вуглеводневих покладів.

Ключові слова: кейси, контрольовані параметри, набуття знань, функція належності, формальні оцінки кейсів, предметна область, експертні оцінки, база кейсів, інтелектуальна система.

Разработана кейс-базирующаяся методика получения знаний при имплементации диалога «эксперт-интеллектуальная система». Предлагается стратегия темпорального представления добытых кейсов данных и знаний о предметной области. Разработана методика инициализации функции принадлежности по методу формальных оценок кейсов, использующаяся для обработки результатов диалога с группой специалистов по контенту. Применен разработанный формально-логический аппарат к обработке кейсов реальных данных предметной области. Проанализирован и обобщен процесс прогнозирования интеллектуальной системой углеводородных запасов. Предлагается результат практической реализации архитектуры базы кейсов интеллектуальной системы для прогнозирования углеводородных залежей.

Ключевые слова: кейсы, контролируемые параметры, приобретение знаний, функция принадлежности, формальные оценки кейсов, предметная область, экспертные оценки, база кейсов, интеллектуальная система.

The case-based method of acquiring knowledge during the implementation of the dialogue "expert-intellectual system" was developed. The strategy of temporal representation of the extracted case data and knowledge about the subject area is proposed. A method for initializing the membership function based on the method of formal case evaluations is developed, which is used to process the results of a dialogue with a group of content specialists. A formal logic has been developed to handle case data of real domain data. The process of prediction of hydrocarbon reserves by the intellectual system is analyzed and summarized. The practical implementation of the architecture of the case base of the intellectual system for prediction of hydrocarbon deposits is proposed.

Key words: cases, controlled parameters, acquisition of knowledges, membership function, formal evaluation of cases, subject area, expert evaluations, case base, intellectual system.

Вступ

Однією із основних цілей побудови інтелектуальних систем є забезпечення можливості ітеративних консультацій, що починаються з простих і легкодоступних даних. Як правило, на основі цих даних може бути виведено число можливих прогнозів, відтак для підтвердження

або спростування цих гіпотез потрібно подальше дослідження. Система використовує нафтогазову інформацію, що зберігається в базі даних, для того, щоб спрогнозувати форми, яких повинні набуті ці додаткові дослідження. Очевидно, що для подальшого дослідження виконують саме ті спостереження, які повинні підтвердити або спростувати певний прогноз. Крім

того, вказуються ті спостереження, які можуть перевести даний очікуваний, ймовірний прогноз на більш високий рівень в ранжованому списку всіх можливих прогнозів.

Мета роботи виділити в предметній області ключові поняття і об'єкти, описати взаємозв'язки між ними і визначити процедури обробки знань. Основна проблема в дискретності і низькій структурованості знань експерта і тих труднощах, які виникають при спробі експерта дати словесний опис тих знань, які він використав для прийняття того чи іншого рішення.

Аналіз сучасних закордонних і вітчизняних досліджень і публікацій

В інженерії знань виділяють дві ключові фігури – експерт та інженер-когнітолог, який проводить бесіди з експертом [1, 2]. Відомо, що доступна експерту інформація про нафтогазові об'єкти і взаємозв'язки між ними в загальному випадку є суттєво невизначеною [3, 4, 5]. Проте експерт все ж здатний робити певні логічні висновки на основі цієї інформації. Л. Заде [6] запропонував оцінювати ступінь належності числам із інтервалу [0..1]. Фіксування конкретних значень при цьому має суб'єктивний характер. Важливим є як характер змін, так і тип шкали, в якій одержують інформацію від експерта. Під час проектування інформаційної системи адаптовано і модифіковано метод інтервальних оцінок, запропонований в роботі [7]. Згідно з теорією конструктивів [8] у свідомості експерта існують суб'єктивні утворення, що описують альтернативні події, протилежні моделі поведінки в певних ситуаціях. Наприклад, колектор нафти і газу – гірська порода, що здатна вміщувати нафту і газ і віддавати їх (пропускати крізь себе). Породо-покришка, на відміну від колектора, не вміщує нафти і газу. Інформація про відношення може бути зібрана в числовому або лінгвістичному вигляді з використанням наперед визначених лінгвістичних значень для означення *контрольованих параметрів*, таких як частота проявів і ступінь підтвердження. Таким чином, можуть бути одержані емпіричні знання, що спираються на міркування, які визначають знання [9, 10].

Висвітлення невирішених раніше частин загальної проблеми

Введення ключових слів викликає відповідні розділи із довідника спостережень. Інженер-когнітолог може зв'язувати з цими спостереженнями відповідні нечіткі значення. Після того, як зібрані всі спостереження, роблять логічний висновок типу “спостереження-спостереження”. Список спостережень містить всі необхідні пункти даних, що включають нечіткі значення, результати вимірювання і ви-

сновку, числові значення, дані спостережень. Після цього список спостережень перевіряється на суперечливість. Підтверджені і будь-які інші можливі прогнози в сукупності повинні пояснювати будь-яке спостереження чи результат лабораторного дослідження. Наявність непояснених даних вказує як правило, на те, що певну сукупність об'єктів необхідно дослідити додатково. *Таким чином*, знання про реальний світ характеризуються неповнотою, неточністю і неузгодженістю, що вимагає застосування теорії нечітких множин, яка уможливує означення нечітких нафтогазових об'єктів і закладає, відповідно, основу лінгвістичного підходу-як ефективного наближення модельованої сутності нафтогазових об'єктів.

Формулювання цілей статті

Однією із основних властивостей якісної консультативної експертної системи є діалог, що оперує добре зрозумілими користувачу поняттями. Так, якщо з системою працює спеціаліст, використання складних термінів робить діалог точним і ефективним, коли діалог проводиться з непрофесіоналом, складні терміни тільки заплутують його. Ці дві групи користувачів представляють границі загальної ситуації створення діалогу для експертної системи.

Таким чином, *метою даної статті* є ймовірна оцінка та інтерпретація даних контрольних параметрів предметної області для спеціаліста-експерта, що дозволяє сформулювати прогноз найбільшої точності, яка залежить від здатності системи пояснити одержані нею результати. За запитом користувача повинна надаватися інформація, що додатково пояснює підтверджені і можливі прогнози. Ця інформація повинна відображатися у вигляді назв нафтогазових об'єктів, їх означень, відповідних їм вимірних нечітких значень *контрольованих параметрів*. Дуже часто користувач відповідає *невизначено* на питання системи, сподіваючись на краще розуміння сутності проблеми у ході діалогу. Дана проблема вирішується шляхом впорядкування діалогу за рівнями. Кожен рівень розрахований на певного користувача. На найвищому рівні використовують складні і точні терміни. Зі зниженням рівня мова діалогу все більше наближається до рівня ключових термінів.

Висвітлення основного матеріалу дослідження

Точність існує тільки як абстракція. Абстракція ж може бути означеною, як здатність людини розпізнавати і вибирати потрібні властивості явищ і об'єктів реального світу. Це створює можливість для побудови концептуальних моделей, що визначають абстрактні класи явищ і об'єктів. Проте, насправді, кожне

явище, кожний об'єкт реального світу є унікальним [11, 12].

Абстрактні моделі явищ і об'єктів реального світу, такі як математичні структури, рівності і висловлювання, є штучними конструкціями. Це ідеальні структури, ідеальні рівності, ідеальні висловлювання.

Для ефективної роботи з експертом інженер зі знань повинен виділити і дослідити всю множину особистісних конструктивів експерта і їх структуру.

Інженер зі знань разом з експертом складають список об'єктів, що відіграють роль гіпотез і спостережень. Наприклад, для аналізу спостереження "приплив нафти і газу" проводять:

- відбір керну (зразків порід);
- геофізичні дослідження свердловини (каротаж);

- випробовування свердловин.

Експерт вказує, які пари «спостереження-гіпотеза» і «гіпотеза-гіпотеза» асоціативно пов'язані. Наприклад:

- чуттєво-візуальні – запах нафти;
- форма кривих при каротажі свердловини і значення амплітуд ;
- одержання припливів нафти і газу.

В результаті одержують сітку об'єктів. На наступному кроці одержують інформацію про тип об'єкта (чи він одержаний в результаті спостереження, чи в результаті висновку), тип асоціативного зв'язку (пояснення, передбачення), напрям і числова оцінка асоціативного зв'язку. Із всіх гіпотез, що пояснюють дане спостереження, вибирають найбільш ймовірну. Наприклад, маючи спочатку дві основні гіпотези «нафтоносність» і «водоносність» результаті за результатами аналізу спостережень і інтерпретації одержаних даних робиться висновок про характер насичення колектора (здебільшого за каротажними кривими).

Далі виділяються окремі фрагменти інформації, і встановлюються зв'язки між кожним фрагментом і текстом, що його уточнює:

- склад колектора – порода (пісок, глина, алевроліт);
- форма каротажних кривих, зв'язок між ними ;
- наявність певного виду флюїду.

Також встановлюються зв'язки між парами фрагментів. Наприклад, невідповідність характеру каротажних кривих гіпотезі про насичення колектора певним флюїдом вказує на наявність винятків.

Для кожної з гіпотез, що пояснюють дане спостереження, вводять певну числову оцінку – середньостатичні дані по району дослідження.

Виділяється ряд фільтраційно-ємнісних властивостей колектора:

- пористість (діапазон значень 8-25 %);
- проникність (діапазон значень 10-1200 мД);

- водонасиченість (діапазон значень 5-100 %);

- нафтонасиченість (діапазон значень 0-95 %);

- газонасиченість (діапазон значень 0-95 %).

В множині всеможливих станів виділяють продуктивні і непродуктивні колектори. Серед продуктивних виділяють нафто- і газонасичені.

Між фрагментами прослідковують такі типи зв'язків: "склад породи – форма числові значення кривих", "товщина породи – форма кривої", "насиченість породи – форма кривої", "насиченість - опір".

На основі одержаної інформації інженер зі знань формує грубу моделі предметної області, означуючи предикати і їх можливі аргументи. Наприклад, предикат – порода, аргументи – пісковик, вапняк, алевроліт, глина; предикат – колектор, аргументи – пісковик, алевроліт, вапняк; предикат – покришка, аргументи – сіль, глина, вапняки.

Властивості предикатів і зв'язки між ними утворюватимуть структуру предметної області. Властивості предикатів і зв'язки між ними розглядають з точки зору взаємного розташування в розрізі земної кори колекторів і покришок. Взаєморозміщення типу "колектор - низу, покришка - зверху" є сприятливим для нафтогазонакопичення.

Далі з допомогою метазнань здійснюється перевірка відповідності між фактами і предикатами, індуктивний висновок фактів із правил, висновок правил із інших правил. В побудованій моделі при достатній кількості структурних компонентів можна виконувати передбачення невідомих фактів і знань шляхом перенесення їх з одних об'єктів на інші, з огляду на подібність між цими об'єктами.

Обов'язковим є виконання таких умов:

- 1) відповідність одержаних каротажних кривих теоретичним положенням;

- 2) відповідність лабораторних досліджень керну результатам випробовувань;

- 3) відповідність висновків про продуктивність ГДС результатам випробовування.

При проведенні досліджень доцільно використовувати аналогії між характером каротажних кривих в одних геологічних умовах із відповідними кривими в інших геологічних умовах.

У процесі побудови моделі для позначення об'єкту найнижчого рівня використовують індивідуальні імена (наприклад, пісковик, вапняк, алевроліт). Загальні імена використовують для позначення невизначеного об'єкту, або множини об'єктів (наприклад, порода, колектор). Для представлення предикатів використовують метаімена. Їм не відповідають ніякі множини об'єктів.

Імена визначають об'єм понять та ознаки, що характеризують їх зміст [13, 14]. При про-

гнозуванні нафтогазових колекторів пропонується застосовувати такі ознаки.

1. Дименесив (характеристика розміру, обчислювальної міри величини) – застосовується для опису понять «пористість», «проникність».

2. Квалітатив (якісні характеристики) – використовуються при описі складу колектора (теригенний, карбонатний) та характеру насичення.

3. Локатив (характеризує місцезнаходження) – використовується для опису взаєморозміщення (наприклад, “колектор-покришка”).

4. Темпоратив (опис часових характеристик) – при розкритті нафтонасиченого колектора і при промисловому відборі нафти відбувається її заміщення водою.

5. Фабрикатив (фізичний склад) – вуглеводні в колекторі можуть перебувати в двох фазових станах: рідина або газ.

6. Генеративний зв'язок Gen (коли один компонент позначає предмет, що належить до сукупності, позначається іншим компонентом) – компонент-зерно, компонент-пора, тріщина. Якщо порода є теригенною і має зернисту структуру, то наявний другий компонент - пора.

7. Дестинативний зв'язок Des (один компонент, який вказує на призначення іншого компонента) – в системі «природний резервуар нафти і газу» є два компоненти: порода-колектор і порода-флюїдоупор. Другий компонент призначений для створення екрану для першого компонента.

8. Директивний зв'язок Dir (один компонент позначає шлях, напрям іншого компоненту) – система природний резервуар зумовлює переміщення флюїдів в напрямі знизу угору.

9. Інструментальний зв'язок Ins (один компонент позначає знаряддя дії, що позначається іншим компонентом) в процесі переміщення флюїдів у колекторі відбувається їх диференціація і розміщення за густинами.

10. Каузальний зв'язок Sous (один компонент позначає причину виникнення іншого компонента через певний проміжок часу) – в процесі розробки родовища методом заводнення пластів закачування води в нагнітальну свердловину має наслідком витіснення нафти з пласту.

11. Комітативний зв'язок Com (один компонент позначає дію, що супроводжує інший компонент) – в процесі розробки нафтового покладу при падінні пластового тиску розчинений в нафті газ, розширюючись, спричиняє витіснення нафти.

12. Корелятивний зв'язок Cor (один компонент виражає можливість спостереження за іншим компонентом, або відповідності одного предмета іншому) – пара «пористість-проникність». Між пористістю і проникністю часто спостерігається прямий зв'язок – збіль-

шення пористості призводить до збільшення проникності.

13. Негативний зв'язок Neg (один компонент заперечує можливість появи іншого компонента) – якщо пластові води супроводжують нафтовий поклад, то вони, як правило, високо-мінералізовані. Наявність прісної води заперечує наявність покладу.

14. Лімітативний зв'язок Lim (один компонент позначає сферу застосування чи призначення іншого компонента) – характер насичення колекторів мінералізованими водами хлор-кальцієвого типу визначає застійний характер гідродинамічного режиму цих вод.

15. Медіативний зв'язок Med (один компонент має значення способу дії другого) – закачування води – витіснення нафти.

16. Посесивний зв'язок Pos (один компонент виражає приналежність іншому компоненту) – у нафті завжди є розчинений вуглеводневий газ.

17. Потенсивний зв'язок Pot (один компонент призводить до можливості виникнення іншого через деякий проміжок часу) – при певних термобаричних умовах (висока температура і тиск) рідинні вуглеводні перетворюються в газоподібні.

18. Результативний зв'язок Res (один компонент виражає наслідок дії іншого) – розширення пластового газу внаслідок зменшення пластового тиску.

19. Репродуктивний зв'язок Rep (один компонент є вихідним пунктом для відтворення чи перетворення іншого компонента) – при підйомі на поверхню газової суміші з великих глибин виділяються два компоненти – газ і конденсат.

20. Ситуативний зв'язок Sit (один компонент позначає ситуацію, що визначає стан або область дії другого) – зменшення пластового тиску нижче тиску насичення спричиняє виділення газу в окрему фазу.

21. Трансгресивний зв'язок Trg (один компонент позначає результат перетворення іншого) – сума водонасичення, газо- і нафтонасичення дорівнює 100%.

Одним із способів прискорення пошуку релевантних знань є їх зв'язаність [4]. Знання слід організовувати навколо найбільш важливих об'єктів предметної області. Всі знання, що характеризують деяку сутність, зв'язуються і подаються вигляді окремого об'єкту.

Якщо системі потрібна інформація про деяку сутність, то вона спочатку шукає об'єкт, що її описує, а потім вже всередині об'єкту шукає інформацію про дану сутність.

В об'єктах, в свою чергу, виділяють два типи зв'язків між елементами – зовнішні і внутрішні. Внутрішні зв'язки об'єднують елементи в єдиний об'єкт і утворюють його структуру. Зовнішні зв'язки відображають залежності, що існують між об'єктами в предметній області.

Зовнішні зв'язки поділяють на логічні і асоціативні. Логічні зв'язки виражають семантичні відношення між елементами знань. Асоціативні зв'язки забезпечують взаємозв'язки, що прискорюють пошук релевантних знань.

Кожен виділений об'єкт предметної області володіє певними характеристиками та властивостями. Наприклад, для об'єкту "родовище" типовими атрибутами є глибина залягання, вік, літологія тощо. Атрибути набувають конкретних значень. Так, глибина залягання може набувати значень 5096 м, літологія – пісковики кварцові тощо

В таблиці 1 наведено опис об'єкту "родовище" [15].

Таблиця 1 – Представлення фактів з допомогою типлетів O - A - 3

Об'єкт	Атрибут	Значення
Родовище	назва	Березівське
Родовище	глибина залягання	5096
Родовище	літологія	пісковики кварцові

Використання триплетів O-A-3 для опису реальних об'єктів не враховує власне характер і вид взаємозв'язків, які можуть мати причинно-наслідковий характер, чи бути відношеннями типу «частина-ціле» тощо. На форму зв'язків не встановлюється жодних обмежень.

Властивості конкретного об'єкта описують парами «атрибут-значення» A-3 (табл. 2).

Таблиця 2 – Представлення фактів з допомогою пар A - 3

Атрибут	Значення
назва	Березівське
глибина залягання	5096
літологія	пісковики кварцові

Відношення між окремими характеристиками об'єктів подається у вигляді таблиць, в яких кількість стовбців відповідає ступеню відношення, а один рядок дорівнює відповідному кортежу.

Дані про кожний об'єкт, що описується в таблиці (про кожне родовище), займають один цілий рядок. Структура всіх рядків збігається. Степінь відношення n вказує на те, що описуються однотипні об'єкти, які мають по n атрибутів. Якщо система зустрине в базі знань відношення A і B з степенями відповідно 5 і 7, то вона розуміє, що йдеться про різнотипові об'єкти, один з яких має 5, а другий 7 атрибутів.

Таблиця 4 – Опис об'єкту "порода"

Порода	Характеристика X	Характеристика Y
Продуктивний колектор	опір_пласту >= 20	Пористість_відкрита>6. 5
Колектор	опір_пласту < 20	Пористість_відкрита>6. 5

В таблиці 3 зображено відношення, куди записані відомості про кілька родовищ, по яких працюватиме система.

Таблиця 3 – Фрагмент відношення степеня 4

Назва родовища	Глибина залягання	Вік	Літологія
Кременівське	1210	C2	пісковики поліміктові
Богатойське1	2232	C2	пісковики мезоміктові
Богатойське3	4297	C1t	вапняки і доломіти
...

Співвідношення родовищ – це одна із іменованих таблиць. Ім'я таблиці (в даному випадку «Родовище») визначаються смисловим відношенням. Степінь відношення дорівнює кількості стовпців у таблиці.

Як ефективні засоби збереження фактів використовуються правила зі змінними величинами. Змінні в правилах замінюють певну сукупність фактів. Розглянемо правила в *формі обмежень*:

обмеження

опір_пласту >= 20
і пористість_відкрита>6. 5

висновок

порода - продуктивний колектор.

обмеження

опір_пласту < 20
і пористість_відкрита>6. 5

висновок

порода - колектор.

Через використання змінних ці правила можна об'єднати в одне:

обмеження

порода має характеристику X
і порода має характеристику Y

висновок

порода - Z.

Для правил із змінними будують реляційні таблиці, в яких містяться фактичні значення змінних.

В таблиці 4 описується вся та інформація, що і в вихідних правилах. Назва таблиці відповідає назві абстрактного об'єкту, а заголовки стовбців – атрибутам об'єкту.

Використання фреймів дозволяє об'єднувати всі характеристики окремого об'єкту в одну групу і маніпулювати з нею, як з єдиним цілим.

Фрейму присвоюється ім'я, що збігається з назвою об'єкту, який він описує. Фрейм складається зі слотів. Вся наявна інформація про об'єкт структурується, і її окремі компоненти записуються в слоти.

В таблиці 5, як приклад, показано фрейм, з допомогою якого описується абстрактна категорія «родовище».

Таблиця 5 – Фрейм «родовище»

Назва родовища	?
Глибина залагання	?
Вік	?
Літологія	?
Дослідження	ДОСЛІДЖЕННЯ
Ознака продуктивного колектора	Продуктивний, якщо опір_пласту > 20

Фрейм містить інформацію, загальну для всіх родовищ. Символ «знак питання» означає, що вміст слоту ще невідомий системі.

Верхня частина фрейму призначена для зберігання фактичної інформації. Відповідні слоти можна було би без проблем записати у вигляді триплетів O - A - Z чи n-арних відношень.

Вмістом слоту «дослідження» посилання на інший фрейм.

В нафтогазовій справі немає потреби розглядати мікроявища і мікрооб'єкти, щоб зіткнутися з проблемами неповноти, невизначеності і неузгодженості. Недостатня кількість інформації, її неточність і суперечливість є дуже важливим фактом.

Розглядаються чотири класи об'єктів:

- спостереження, ознаки, результати геофізичних досліджень і одержані дані (tp_i);
- гіпотези і сценарії (hyp_j);
- проміжні комбінації ($temp_k$);
- комбінації спостережень ($\otimes tp_l$).

Спостереження tp_i приймають значення mf_{OVi} із $[0,1] \cup \{v\}$. Значення mf_{OVi} вказує на інтенсивність прояву спостереження tp_i . Після цього будується бінарне нечітке відношення $W_{pOB} \subset \Pi \times \Sigma$, що визначається умовою $mf_{Wdeob}(DEq, tp_i) = mf_{OVi}$ для родовища DEq, де $DEq \in \Pi$ ($\Pi = \{DE1, \dots, DEr\}$) і $tp_i \in \Sigma$ ($\Sigma = \{tp_1, \dots, tp_m\}$).

Прогнози також набувають значень із $[0, 1] \cup \{v\}$. Нечіткі значення $0.00 < mf_{hyp_i} < 1.00$ – це можливі прогнози, тоді як значення $mf_{hyp_i} = 1.00$ і $mf_{hyp_j} = 0.00$ відповідають підтверженому і непідтверженому прогнозам від-

повідно. Ще не розглянутим прогнозам призначають значення $mf_{hyp_j} = v$. Формально будується співвідношення $W_{DEFC} \subset \Pi \times \Delta$, що визначається умовою $mf_{Wdefc}(DEq, hyp_j) = mf_{hyp_j}$ для родовища DEq, де $hyp_j \in \Delta$ ($\Delta = \{hyp_1, \dots, hyp_n\}$).

Обидва об'єкти, що розглядаються: проміжні комбінації і комбінації спостережень – набувають значень mf_{fck} і mf_{OCI} відповідно із $[0, 1] \cup \{v\}$, де v означає, що дійсне значення ще не визначено. Відношення $R_{DEOVC} \subset \Pi \times K$ задається умовою $mf_{deobc}(DEq, \otimes tp_l) = mf_{OVC}$ для родовища DEq, де виконується $OCI \in K$ ($K = \{\otimes tp_1, \dots, \otimes tp_l\}$), і W_{DEOVC} формально описує комбінації спостережень по даному родовищу.

Нечіткі логічні зв'язки визначаються наступним чином:

$Goal_1 \wedge Goal_2 = \min\{Goal_1, Goal_2\}$, якщо $Goal_1 \in [0, 1]$ і $Goal_2 \in [0, 1]$;

$Goal_1 \wedge Goal_2 = v$, якщо $Goal_1 = v$ і/або $Goal_2 = v$;

$Goal_1 \vee Goal_2 = \max\{Goal_1, Goal_2\}$, якщо $Goal_1 \in [0, 1]$ і $Goal_2 \in [0, 1]$;

$Goal_1 \vee Goal_2 = Goal_1$, якщо $Goal_1 \in [0, 1]$ і $Goal_2 = v$;

$Goal_1 \vee Goal_2 = Goal_2$, якщо $Goal_1 = v$ і $Goal_2 \in [0, 1]$;

$Goal_1 \vee Goal_2 = v$, якщо $Goal_1 = v$ і $Goal_2 = v$;

$Goal_1 = 1 - Goal_1$, якщо $Goal_1 \in [0, 1]$;

$Goal_1 = v$, якщо $Goal_1 = v$.

Слід зазначити, що введення величини v для відсутніх операндів, яка обов'язково повинна бути включена в означення «зв'язок і заперечення» із практичних міркувань, призводить до порушення деяких законів (наприклад, законів де Моргана). Ці закони справедливі для класичних нечітких зв'язок

$$Goal_1 \wedge Goal_2 = \min\{Goal_1, Goal_2\},$$

$$Goal_1 \vee Goal_2 = \max\{Goal_1, Goal_2\}$$

відносно заперечення $Goal_1 = 1 - Goal_1$, але тепер вони втрачають силу.

Між нафтогазовими об'єктами розглядаються такі відношення:

- «спостереження-прогноз» ($tp_i hyp_j$);

- «комбінація спостережень – прогнозів» ($\otimes tp_l hyp_j$);

- «спостереження – спостереження» ($tp_i tp_j$);

- «прогноз-прогноз» ($hyp_i hyp_j$).

Ці відношення характеризуються двома параметрами:

- частота прояву (f);

- ступінь підтвердження (cd).

Відношення між нафтогазовими об'єктами задаються у вигляді правил відношення з відповідними їм парами відношень. Загальне формулювання такого правила має вигляд:

Таблиця 6 – Лінгвістичні нечіткі значення

Частота прояву			Ступінь підтвердження		
Значення λ_f	Інтервал	Характерне значення mf_f	Значення λ_{cd}	Інтервал	Характерне значення mf_{cd}
завжди	[1.00, 1.00]	1.00	завжди	[1.00, 1.00]	1.00
майже завжди	[0.98, 0.99]	0.99	майже завжди	[0.98, 0.99]	0.99
дуже часто	[0.83, 0.97]	0.90	дуже сильно	[0.83, 0.97]	0.90
часто	[0.68, 0.82]	0.75	сильно	[0.68, 0.82]	0.75
середнє	[0.33, 0.67]	0.50	середнє	[0.33, 0.67]	0.50
зрідка	[0.18, 0.32]	0.25	слабко	[0.18, 0.32]	0.25
дуже рідко	[0.03, 0.17]	0.10	дуже слабо	[0.03, 0.17]	0.10
майже ніколи	[0.01, 0.02]	0.01	майже ніколи	[0.01, 0.02]	0.01
ніколи	[0.00, 0.00]	0.00	ніколи	[0.00, 0.00]	0.00
невідомо		v	невідомо		v

Примітки: А – пористість відкрита, В – абсолютна проникність, С – опір пласту, D – питома продуктивність, Е – коефіцієнт водонасичення.

ЯКЩО (антецедент) **ТО** (консеквент) **ПРИ** (ОБМЕЖЕННЯ)

Пари відношення (f, cd) містять або числові значення mf_f і mf_{cd} , або лінгвістично нечіткі значення λ_f і λ_{cd} , або і t_i , і інші. Важливість характерних значень полягає в тому, що вони роблять нечіткий висновок легкоздійснюваним. Нижче наведені деякі приклади правил відношень.

Приклад 1

ЯКЩО (пористість_відкрита ≥ 8)
ТО (порода – колектор)

ОБМЕЖЕННЯ (0.75 = часто, 0.25 = слабо)

Приклад 2

ЯКЩО (абсолютна_проникність ≥ 0.7)
ТО (порода – продуктивний_колектор)

ОБМЕЖЕННЯ (0.25 = рідко, 1.00 = завжди)

Приклад 3

ЯКЩО (абсолютно_вільний_дебіт ≥ 20 \wedge питома_продуктивність ≥ 0.006 \wedge опір_пласту ≥ 10 \wedge коефіцієнт_водонасичення ≤ 42 \wedge параметр_насиченості ≥ 5.5)

ТО (порода – продуктивний_колектор)

ОБМЕЖЕННЯ (v, 0.90 = дуже сильно)

Значення mf_f і mf_{cd} інтерпретуються як значення нечітких відношень між антецедентами і консеквентами. Так,

hyp_j (відношення прояву)

$$W_{ОВFC}^f \subset \Sigma \times \Delta$$

$tp_i hyp_j$ (відношення підтвердження)

$$W_{ОВFC}^{cd} \subset \Sigma \times \Delta$$

$\otimes tp_i hyp_j$ (відношення прояву)

$$W_{ОВFC}^f \subset K \times \Delta$$

$\otimes tp_i hyp_j$ (відношення підтвердження)

$$W_{ОВFC}^{cd} \subset K \times \Delta$$

$tp_i tp_j$ (відношення прояву)

$$W_{ОВОВ}^f \subset \Sigma \times \Sigma$$

$tp_i tp_j$ (відношення підтвердження)

$$R_{ОВОВ}^{cd} \subset \Sigma \times \Sigma$$

$hyp_i hyp_j$ (відношення прояву)

$$R_{FCFC}^f \subset \Delta \times \Delta$$

$hyp_i hyp_j$ (відношення підтвердження)

$$R_{FCFC}^{cd} \subset \Delta \times \Delta$$

Система набуття знань здатна одержувати інформацію про нафтогазові об'єкти і відношення між ними. Відношення зберігаються у вигляді числових значень із проміжку [0, 1]. Інформація може бути одержана одним із двох способів:

- з допомогою числових або лінгвістичних оцінок від спеціалістів-експертів;

- шляхом статистичного оцінювання бази даних, що містить інформацію про нафтогазові об'єкти з підтвердженими результатами прогнозів.

Важливою властивістю таких відношення є те, що вони можуть бути інтерпретовані статистично. Частота прояву f і ступінь підтвердження cd становить:

$$mf_f = G(tp_i \cap hyp_j) / G(hyp_j) = G(tp_i / hyp_j) \quad (1)$$

$$mf_{cd} = G(tp_i \cap hyp_j) / G(tp_i) = G(hyp_j / tp_i) \quad (2)$$

де $G(tp_i \cap hyp_j)$ - абсолютна частота прояву tp_i і hyp_j ;

$G(hyp_j)$ - абсолютна частота прояву hyp_j ;

$G(tp_i)$ - абсолютна частота прояву tp_i ;

$G(tp_i / hyp_j)$ - умовна частота tp_i за умови hyp_j ;

$G(hyp_j / tp_i)$ - умовна частота hyp_j за умови tp_i .

З допомогою означень (1) і (2) розширене статистичне оцінювання відомих відношень між нафтогазовими об'єктами, або ще не ви-

Таблиця 7 – Зведена таблиця числового і лінгвістичного оцінювань родовищ

Родовище	A	B	C	D	E
Атазівське	7.0	досить низьке	досить високе	невідомо	[40-43]
Березівське	близько 8.0	[0.7-0.9]	високе	низьке	41
Буріватівське	[6.9- 8.1]	0.8	[9-10]	середнє	близько 42
Західносолохівське	низьке	досить високе	дуже високий	[0.005- 0.013]	[41-44]
Котелевське	близько 7.6	близько 1	низьке	дуже високе	досить високе
Матвійвське	середня	[0.7-0.8]	9.5	високе	низьке
Новомиколаївське	[7.9- 8.7]	Невідомо	близько 9	0.0075	середнє

значених відношень може бути здійснено використанням даних про вже розвідані і вивчені родовища.

В таблиці 7 наявні декілька типів значень атрибута: точні значення, інтервальні значення, нечіткі значення (наприклад, “високе”, “близько 9”, “середнє”) і невизначені значення. Неточні значення атрибута відповідають або частковому (неповному, або нечіткому) знанню, або грубою оцінкою.

В таблиці 8 область значень атрибута ТИП дискретна: $D_{\text{тип}} = \{a, b, c, d\}$, де кожне число відповідає певному типу характеристик. Оскільки тут наявні області двох типів (неперервні і дискретні), то можливі два способи подання функцій розподілу можливостей.

Таблиця 8 – Відношення з кількома атрибутами

Характеристика1	Характеристика2	Тип
пористість	проникність	розмір
колектор	покришка	місцезнаходження
рідина	газ	фізичний склад
закачування води	витіснення нафти	спосіб дії

Припустивши, що ступені приналежності нам відомі, наприклад у вигляді $mf_{\text{OB}}(v_i) = \xi_i$ ($i=1, 2, \dots, m$), то попарні порівняння можна представити матрицею відношень $C=(c_{ij})$, $c_{ij}=\xi_i/\xi_j$.

Оскільки відношення порівняння в реальному світі неточні через свою емпіричність, то потрібно обчислювати окремі оцінки для ξ . В процесі розв’язання задачі формується матриця порівнянь розглядуваної множини елементів. Елементи матриці показують, наскільки один елемент кращий від іншого. При формуванні оцінок попарних порівнянь від експерта вимагається вираження досвіду у вигляді:

- встановити, який із двох даних елементів є більш важливим;

- оцінити сприйняття інтенсивності розрізнення у вигляді рангу важливості за відповідною шкалою.

Розглядаючи поняття “клас F”, яке описується функцією належності на множині об’єктів $C=\{c^0, \dots, c^{m-1}\}$. В C є тільки два об’єкти c^k та c^l , з яких c^k - ідеальний представник тих об’єктів, які належать C, а c^l - ідеальний представник тих об’єктів, що не належать цьому поняттю. Завдання експерта полягає в ранжуванні об’єктів відповідно до їх належності (чи неналежності) описаному поняттю. Результируюча матриця попарних порівнянь задає порядок пар об’єктів за ступенями відмінностей в парах.

На другому етапі визначається та максимальна кількість класів, яка може бути описана даним набором параметрів. Для кожного елемента v значення функції належності класу F_1 доповнює до одиниці значення функції належності класу F_2 . Таким чином, система класів повинна складатися із класів, що описують протилежні події. Сума значень функції належності довільного елемента v до системи таких класів буде дорівнювати одиниці. У випадку, коли число класів і їх склад чітко невизначений, то вводиться додатковий віртуальний клас, що включає всі невиявлені класи. Завдання експертів полягає в процентній оцінці ступеня прояву кожного з класів у конкретному стані v .

Вибірку об’єктів доцільно брати такою, щоб рівномірно представити ступені належності від 0 до 1 відносно нечіткої множини, що розглядається. При цьому повинні бути присутні принаймні два об’єкти, ступені належності яких рівні відповідно 0 і 1. В результаті опитування експертів одержується процентна оцінка ступенів належності.

Нехай інтервал $[y_{ml}, y'_{ml}]$ відображає думку l -го експерта ($l=1, \dots, n_1$) про значення m -ї ($m=1, \dots, n_2$) ознаки оцінюваного поняття ОВ. Повний опис цього поняття l -им експертом подаватиметься у вигляді

$$\Omega_l = [y_{1l}, y'_{1l}] \times \dots \times [y_{n_2l}, y'_{n_2l}] \quad (3)$$

Наведений нижче алгоритм, що дозволяє обчислювати коефіцієнти компетентності експертів, а також приводити вихідну “розміту” функцію, що є середнім значенням експертних оцінок, до характеристичної функції нерозмитої чіткої множини:

1. Для кожної ознаки m розглядають всі запропоновані експертами інтервали, знаходять їх об’єднання, які складається з інтервалів, що не перетинаються:

$$[y_{mk}, y'_{mk}], (m=1, \dots, n1 ; k=1, \dots, n2) . \quad (4)$$

2. На основі одержаних множин будемо результуючу множину $Goal_k$:

$$Goal_k = [y_{1k}, y'_{1k}] \times \dots \times [y_{nk}, y'_{nk}], k=1, \dots, n2 \quad (5)$$

3. Обчислюємо для $x \in Goal_k$

$$\xi_l(y) = \begin{cases} 1, Z_k \cap \Omega_l \neq \emptyset \\ 0, Z_k \cap \Omega_l = \emptyset \end{cases} . \quad (6)$$

4. Приймаємо номер ітерації $in=1$

5. Вводимо коефіцієнти компетентності для експертів :

$$\{CC_l^{in}\}_{l=1}^{n1} = \{1/n1\}_{l=1}^{n1} . \quad (7)$$

6. Обчислюємо наближення функції належності при нормованих

$$CC_l : \sum CC_l^{ni} = 1, f^{in}(y) = \sum_{l=1}^{n1} \phi_l(y) * CC_l^{in}, \quad (8)$$

$$y \in Z_k .$$

7. Обчислюємо функціонал неузгодженості думки l-го експерта з думкою групи експертів на in – ітерації :

$$FD_l^{in} = \sum_{y \in Z_k} [f^{in}(y) - \phi_l(y)]^2, l=1, \dots, n1 . \quad (9)$$

8. Обчислюємо

$$\Phi = \sum_{l=1}^{n1} 1 / FD_l^{in} . \quad (10)$$

9. Збільшуємо крок ітерації на одиницю.

10. Обчислюємо

$$CC_l^{in} = \Phi / FD_l^{in-1} . \quad (11)$$

11. Якщо тепер величина $\max |CC_l^{in-1} - CC_l^{in}|$ близька до нуля, то ітераційний процес припиняється, і за наближення функції належності приймається $f(y)=mf_s(y)$, в протилежному випадку виконується перехід до пункту 6.

Основним елементом, з яким оперує система, є функція двох змінних, яка задана на регулярній сітці (полі). У вигляді полів подаються геолого-геофізичні ознаки досліджуваної структури і сам геолого-геофізичний прогноз.

Формовані системою геолого-геофізичні ознаки описують як локальні особливості середовища в околах вузлів, так і крупніші структурні неоднорідності. Геолого-геофізичні ознаки вибираються і кодуються спеціалістами так, щоб поле шуканого прогнозу могло бути з достатньою точністю апроксимовано функцією полів цих ознак. Основні засоби системи призначені для того, щоб допомогти спеціалісту знайти і проаналізувати дану прогнозуючу функцію і одержане поле прогнозу.

Для знаходження прогнозуючої функції використовується вибірка пунктів регіону, для якого даються оцінки прогнозу [16, 17]. У вибірці входять випадково вибрані точки, і точки, по яких існує додаткова інформація стосовно прогнозу. Точки вибірки повинні, по можливості, рівномірно заповнювати регіон і мати рівну

представленість для різних значень прогнозованої величини.

Оцінки прогнозу в точках вибірки роблять експерти. Спеціалісти, що приймають участь в експертизі, дають оцінки незалежно один від одного. При оцінюванні враховуються зареєстровані події або розвідані геологічні об'єкти, відомості про прогнозоване явище і знання про особливості геологічної будови і тенденціях розвитку регіону, що досліджується.

Як оцінки можуть використовуватися зареєстровані значення прогнозу. Проте в ряді випадків відмова від експертних оцінок і формальна заміна їх даними про зареєстровані спостереження збіднюють вхідну інформацію і можуть суттєво погіршити результати прогнозування.

Прогнозуюча функція [12] $FC(x, g)$, де x – вектор ознак; g – вектор оцінюваних параметрів, шукається в деякому наперед вибраному класі функцій, виходячи із умов найкращої апроксимації оцінок прогнозу в точках вибірки.

Важливим питанням є оцінювання точності апроксимації залежності прогнозу від геолого-геофізичних ознак. Відомі методи оцінювання точності апроксимації вимагають введення ряду статистичних тверджень, що явно не впливають із задачі, яка розв'язується. Тому як формальний показник точності рішення використовується величина середньої помилки апроксимації оцінок прогнозу.

Величина середньої похибки апроксимації не може служити єдиним критерієм правильності прогнозу. При хорошому прогнозі в середньому можуть існувати окремі зони, в яких прогноз сильно відрізняється від представлень експерта. В результаті аналізу одержаних розбіжностей спеціаліст повинен дати геолого-геофізичну інтерпретацію одержаного прогнозу і прогнозуючої функції, а також прийняти рішення про необхідність корекції моделі прогнозованого явища. При цьому може бути змінений вид прогнозуючої функції, або внесені зміни в склад вихідних даних: введені або знайдені засобами системи додаткові поля ознак, змінене кодування ознак, додані додаткові точки вибірки тощо. На кожному кроці ітерації спеціаліст формує модель прогнозованого явища і готує вхідні дані. Далі знаходиться прогнозуюча функція. Функція інтерпретується як гіпотеза про шукану закономірність, що узгоджується з наявними знаннями і даними. Результати аналізуються, і спеціаліст приймає рішення про необхідність виконання наступної ітерації.

В системі використовуються такі типи даних [18, 19, 20-24]:

1. Каталоги зареєстрованих геологічних об'єктів (розвідані родовища).

2. Лінеаментні і кільцеві структури, вихідні і одержані засобами системи поля геолого-геофізичних ознак.

Знання експертів складаються із знань про регіональну модель прогнозованого явища і знань про величину прогнозу на вибірці пунктів регіону.

Знання про характер залежності між прогнозованою величиною і геолого-геофізичними ознаками мають якісний характер. Із загальних теоретичних положень, або користуючись якісними моделями і емпіричними даними, спеціаліст може визначити характер зміни прогнозованої величини в залежності від змін окремо кожної із ознак за умови незмінності решти. Ці знання використовуються для вибору геолого-геофізичних ознак, для розроблення способів їх кодування, методики роботи з картами, обґрунтування рекомендацій щодо обчислення вторинних ознак засобами системи, а також для знаходження і аналізу прогнозуючої функції.

В системі функція прогнозування шукається у вигляді :

$$\text{Case}(x, g) = \sum_{i=1}^I \psi_i(x_i, g), \quad (12)$$

де $\psi_i(x_i, g)$ – кусково-лінійна функція від ознаки x_i , в якій вузли наперед задані, а вектор параметрів g визначає значення функції в вузлах. Вибраний клас функцій дозволяє при відносно малій розмірності вектора параметрів врахувати наявні в спеціалістів уявлення про неперервність залежності прогнозу від геолого-геофізичних ознак і про характер нелінійності шуканої закономірності.

Крім цього, цей клас функцій зручний при інтерпретації результатів прогнозу: кожен із функцій $\psi_i(x_i)$ можна інтерпретувати як нелінійний внесок ознаки x_i в прогноз, а поле прогнозування – як суму полів нелінійно перетворених ознак.

Частина знань про характер зв'язку між прогнозом і ознаками представлена в системі у вигляді фреймів і використовується для підтримки пояснення отриманих результатів [21].

Знання про значення прогнозу на вибірці пунктів регіону формалізовані і представлені у вигляді інтервальних експертних оцінок.

Вид інтервальної експертної оцінки EE підбирався таким чином, щоб методика оцінювання була простою і зручною для експерта і водночас забезпечувала одержання достатньо повної інформації про значення геофізичного прогнозу :

$$\text{Case}_1 = (tp^{(1)}, tp^{(2)}, w^{(1)}, w^{(2)}), \quad (13)$$

де $p^{(1)}, p^{(2)}$ – границі інтервалу для значення прогнозованої величини. В задачах апроксимації функціональних залежностей інтервал $[p^{(1)}, p^{(2)}]$ вибирається так, щоб всередині інтервалу всі значення прогнозованої величини були би, на думку експерта, найбільш імовірними і рівноможливими. В задачах розпізнавання прогнозуюча функція повинна набувати різних значень за наявності і відсутності корисних копалин. Із цих міркувань вибираються інтер-

вали $[p^{(1)}, p^{(2)}]$ для комірок, де зареєстровані родовища корисних копалин, і для комірок, де, на думку експертів, родовищ бути не може. $w^{(1)}, w^{(2)}$ – вагові коефіцієнти, за допомогою яких експерт вказує міру своєї впевненості в можливості того, що значення прогнозованої величини може бути меншим або більшим відповідної границі інтервалу $p^{(1)}$ або $p^{(2)}$.

Головна математична задача, що вирішується при побудові прогнозу, є задачею апроксимації функції багатьох змінних в деякому класі функцій [118]

$$\{\text{Case}(x, g)/g \in G\}, \quad (14)$$

де G – область допустимих значень вектора параметрів g . Розв'язок полягає в знаходженні оцінки вектора параметрів g за вибіркою спостережень $\{tp_{j,n}, x_n\}$, де $tp_{j,n}$ – інтервальна оцінка значення прогнозованої величини, визначена j -м експертом в n -й точці вибірки; x_n – вектор геолого-геофізичних ознак в n -й точці.

Оцінка вектора параметрів g має вигляд [22]

$$\bar{g} = \arg \min_{g \in G} \sum_n \sum_j pf(FC(x_n, g), tp_{j,n}), \quad (15)$$

де $pf(FC, S)$ – штрафна функція, що визначає величину штрафу за неточність апроксимації експертної оцінки tp значенням прогнозуючої функції

$$pf(FC, tp) = (w^{(1)}(|FC - p^{(1)}| + p^{(1)} - FC)/2 + w^{(2)}(|FC - p^{(2)}| - p^{(2)} + FC)/2)^q, \quad (16)$$

$$w^{(1)} > 0, w^{(2)} > 0, p^{(1)} \leq p^{(2)}, q \geq 1$$

Якщо прогнозуюча функція $FC(x, a)$ є лінійною за параметрами, то мінімізуючий функціонал є опуклим. Якщо при цьому область G допустимих значень вектора g також опукла, то при оцінюванні можна скористатися ітераційними алгоритмами градієнтного типу.

Очевидно, що в частковому випадку, коли використовуються експертні оцінки, в яких завжди $p^{(1)} = p^{(2)}$, $w^{(1)} = w^{(2)}$ при $q=1$, алгоритм оцінювання збігатиметься з методом найменших модулів, а при $q=2$ – з методом найменших квадратів.

В експертних оцінках неминуче зустрічаються протиріччя. Можлива ситуація, коли оцінки великого числа менш компетентних експертів можуть значно вплинути на оцінки найбільш кваліфікованих експертів. Формальна процедура вибору найбільш компетентних експертів базується на незалежності встановлення експертних оцінок і реалізує просту ідею: представлення і експертні оцінки компетентних експертів краще відображають шукані закономірності і тому близькі між собою.

В системі при знаходженні прогнозуючої функції одночасно виконується вибір найбільш суттєвих ознак. Це здійснюється з допомогою покрокової процедури. Вибирається ознака, що дає найкращу апроксимацію спостережень. На наступному кроці до неї добирається друга ознака, яка дає найменшу похибку апроксима-

ції в парі з першою ознакою і т. д.. Починаючи з деякого кроку похибка апроксимації на навчальній вибірці починає спадати незначно, а похибка на контрольній вибірці починає зростати, що вказує на недоцільність подальшого нарощування ознак.

Результатами рішення є побудований прогноз і знайдена залежність прогнозованої величини від геолого-геофізичних ознак. Мета аналізу полягає в тому, щоб обґрунтувати прогноз і прийняти рішення про спосіб його можливої корекції. Для вибраного класу функцій прогноз природним чином розпадається на адитивні компоненти. Це дозволяє пояснювати значення прогнозу в аналізованих точках, використовуючи значення $\psi_i(x_i, a)$ – внесків ознак x_i в прогноз $FC(x, a)$.

Формалізувати всі міркування спеціаліста про характер шуканої залежності наперед не вдається. Спеціаліст намагається інтерпретувати знайдену залежність в термінах якісних представлень. При одержанні неінтерпретованих функцій $\psi_i(x_i)$ спеціаліст може скоригувати залежність: вибрати інші вузли в функції $\psi_i(x_i, g)$, змінити кодування ознаки x_i , або відмовитися від цієї ознаки взагалі. Інтерпретація підтримується візуалізацією функцій $\psi_i(x_i, g)$ і процедурою пояснення прогнозу в точках вибірки. Процедура пояснення дає текстову інтерпретацію впливу на прогноз в аналізованій точці кожної з ознак. Зіставлення апріорних знань про характер функцій $\psi_i(x_i, g)$ зі знайденою залежністю може виявитися підказкою для нової інтерпретації залежності, або може бути прийнята за основу для коригування моделі.

Метою процедури обґрунтування є підтвердження прогнозу в деякій точці або вказання причини його неточності. При обґрунтуванні знаходиться група точок, що подібна до аналізованої за певними ознаками. Для деяких із виявлених точок може бути наявна додаткова інформація про геофізичний прогноз, наприклад зареєстровані значення прогнозу, або експертні оцінки прогнозованої величини. Ця інформація може підтверджувати прогноз в точці, що аналізується, або заперечувати його. Спеціаліст-експерт може з цим погодитися, або ні. В випадку незгоди перед ним виникає конкретна задача: пояснити, чим аналізована точка відрізняється від точок, виділених процедурою. Це може настановити спеціаліста на висновок про необхідність введення додаткових ознак, про необхідність зміни кодування ознак, про додавання точок у вибірку.

Результати спостережень можуть бути введені трьома способами:

- шляхом введення спостережень tp_i з допомогою природної мови;
- шляхом введення з допомогою природної мови ключевих слів, які приводять в дію цілі групи спостережень tp_i ;

- шляхом оцінювання бази даних, що містить інформацію про об'єкт, і перетворення інформації з допомогою нечіткого інтерпретатора.

На наступному кроці оцінюються проміжні комбінації спостережень. Після перевірки на відповідність обчислюються нечіткі значення для всіх комбінацій спостережень. Одержані в результаті списки тепер повні і не містять жодних протиріч. Нечіткі значення $=1.00$, тобто підтвержені прогнози FC_j для родовища DE_q визначаються з допомогою такого співвідношення:

$mf_{hyp_j} = 1.00$, якщо $mf_{w_{DEFC}}^1(DE_q, hyp_i) = 1.00$
або

$$mf_{w_{DEFC}}^4(DE_q, hyp_j) = 1.00. \quad (17)$$

Нечіткі значення $mf_{hyp_j} = 0.00$, тобто виключені прогнози hyp_i для родовища DE_q визначаються з допомогою : $mf_{hyp_j} = 0.00$, якщо

$$(mf_{w_{DEFC}}^2(DE_q, hyp_j) = 1.00 \text{ або}$$

$$mf_{w_{DEFC}}^3(DE_q, hyp_j) = 1.00$$

$$\text{або } mf_{w_{DEFC}}^5(DE_q, hyp_j) = 1.00$$

$$\text{або } mf_{w_{DEFC}}^6(DE_q, hyp_j) = 1.00).$$

Відношення типу “колектор-продуктивний колектор” допускають логічний висновок наступних прогнозів (такі, що підтверджуються або такі, що виключаються):

$$\mu_{hyp_i} = \begin{cases} 1.00, \text{ якщо } mf_{w_{DEFC}}^{10}(DE_q, hyp_i) = 1.00 \\ 0.00, \text{ якщо } mf_{w_{DEFC}}^{11}(DE_q, hyp_i) = 1.00 \\ 0.00, \text{ якщо } mf_{w_{DEFC}}^{12}(DE_q, hyp_i) = 1.00. \end{cases} \quad (18)$$

Можливим прогнозам відповідають нечіткі значення mf_{hyp_j} , такі, що

$$\begin{aligned} \varepsilon \leq mf_{hyp_j} \leq 0.99, \quad mf_{hyp_j} = \\ = \max[mf_{w_{DEFC}}^1(DE_q, hyp_i); mf_{w_{DEFC}}^4(DE_q, hyp_i); \\ mf_{w_{DEFC}}^{10}(DE_q, hyp_i)], \end{aligned}$$

якщо

$$(\varepsilon \leq mf_{w_{DEFC}}^1(DE_q, hyp_i) \leq 0.99 \quad (19)$$

$$\text{або } \varepsilon \leq mf_{w_{DEFC}}^4(DE_q, hyp_i) \leq 0.99$$

$$\text{або } \varepsilon \leq mf_{w_{DEFC}}^{10}(DE_q, hyp_i) \leq 0.99).$$

Оскільки значення μ_{FC_j} не залежать від кількості правил, які можуть бути використані для підтримки hyp_j . Це може бути введена евристична функція, що враховує число критеріїв, присутніх, або частково присутніх, які передбачають, але не підтверджують істинність прогнозу hyp_i . Таким чином, ця функція обчислює відповідне число точок PNR_{hyp_j} . Ці значення корисні тоді, коли треба робити вибір між різними можливими прогнозами, хоча кінцевою метою повинно бути одержання підтвер-

дженого прогнозу. Число точок PNR_{FCj} обчислюється таким чином:

$$PNR_{hyp_j} = 100 \sum_{i=1}^{k^*} \{ \alpha \min[mf_{WDEOB}(DE_q, tp_i); mf_{W_{OBF}^f}(tp_i, hyp_i)] + \beta \min[mf_{WDEOB}(DE_q, tp_i); mf_{W_{OBF}^{cd}}(tp_i, hyp_i)] \}, \quad (20)$$

де k^* - число спостережень по родовищу DE_q , які виникають при означенні hyp_j , а $\alpha + \beta = 1$.

00. Ми приймаємо $\alpha = 0.09$ і $\beta = 0.91$, тобто степінь підтвердження береться в десять разів більш значущою, ніж частота прояву за відношенням до значення PNR_{hyp_j} . Суму перемножуємо на 100 для того, щоб досягти одержання легких для читання і запам'ятовування значень числа точок.

Представлення знань в системі базується на таких пунктах:

Cases (Розділи)

Parameters (Параметри)

Для більшої читабельності інформації можна включати також коментарі.

Всі компоненти бази знань створюються з допомогою спеціальної мови формальних описів, в якій діють єдині правила синтаксису. Наступні метасимволи мови мають спеціальне призначення:

$::=$ інтерпретується як 'складається з' або 'оголошений як';

{ } фігурні дужки вказують на те, що послідовність в дужках може бути повторена нуль або більше разів;

[] квадратні дужки вказують на те, що послідовність в них використовується як опція;

< > імена в кутових дужках є синтаксичними елементами. Всі синтаксичні елементи будуть визначені, тобто вони з'являться зліва від символу ' ::= ' в граматичному правилі;

| вертикальна лінія читається, як 'або' і використовується для розділення альтернатив.

Верхній рівень представлення знань складається з розділів. Перший розділ в будь-якій базі знань повинен мати ім'я Start (Початок). Розділ складається з імені, текстового опису і кількох параграфів. Система опрацює параграфи розділу по одному зверху до низу. Якщо параграф містить логічний вираз, то він обчислюється першочергово. Якщо цей вираз істинний, або параграф не містить логічних виразів взагалі, то всі дії виконуються згідно з порядком, заданим в параграфі.

Синтаксис додатку:

section-name - ім'я розділу. Може складатися з букв, цифр і знаків підкреслення;

description-text - послідовність символів взятих в лапки;

boolean-expression - логічні вирази, або умови - забезпечують логічне підтримання ведення діалогу. Вони використовуються як частини if-операторів в оголошеннях розділів і поля правил параметрів. Логічні вирази можуть

набувати одного із значень - істина, хиба, невізначено. Вони складаються з:

- логічних параметрів;
- числових відношень;
- відношень між символьними іменами і змінними;
- складених логічних виразів.

Відношення між числовими і символьними змінними записуються через використання реляційних операторів (<, <=, =, >, >=, <>).

Символьні рядки порівнюються посимвольно зліва направо. Порівняння символів здійснюється через використання їх ASCII - кодів.

Складені логічні вирази формуються з використанням логічних операторів: and, or і not. Оператор not має найвищий пріоритет. Оператори and і or мають пріоритет нижчий, ніж в математичних операцій. За допомогою дужок можна змінювати порядок виконання операцій.

Функція know повертає значення «істина», якщо параметр набуває певного значення, і «хиба» - в протилежному випадку.

advice - порада, є результатом обчислення значень параметрів, виконання певної послідовності розділів і параграфів.

text-expression - текстові вирази, використовуються в полях пояснення (explanation field), полях запитань (question field) і в правилах. В розділах текстові вирази використовуються як поради (advice). Текстові вирази утворюються, як суміш рядків символів, значень певних параметрів і текстових функцій. Використання імені параметра дозволяє викликати значення цього параметра, що відображається, як частина складеного тексту. Символ & використовується для вставки в текст нового рядка.

assign - присвоїти, використовується для присвоєння параметру певного значення. Синтаксис:

<assign> ::= assign <parameter-name> := <expression>

Тип результату виразу повинен співпадати з типом параметра.

Використання акції assign дозволяє поєднувати можливості прямого виведення (виведення від фактів до мети, forward chaining) розділів з можливостями зворотного виведення (виведення від мети до фактів, backward chaining) в параметричних правилах.

call - використовується для виклику однієї з вбудованих, чи означених користувачем процедур.

call clear_all() - обнулює значення всіх параметрів поточної бази знань;

call clear_value(<parameter_name>) - обнулює значення конкретного параметра;

call display(<filename>) - виводить на екран вміст текстового файлу;

call hyperadvice(<filename>, <name>) - викликає певний розділ Windows-help файлу;

call save_values(<filename>) - записує в файл значення всіх параметрів;

`call restore_values(<filename>)` - надає параметрам тих значень, які були попередньо записані в відповідний data-файл ;

`call system(<string>)` - використовується для запуску з системи інших DOS і Windows прикладних програм ;

`chain` - дія, що дозволяє системі звернутися до нової бази знань. Синтаксис:

```
<chain> ::= chain <filename>
```

Структурування знань з розподілом їх в окремі бази розширює можливості системи обробки інформації. Значення параметрів можуть передаватися від однієї бази знань до іншої через виконання операції `save_values` перед початком логічного висновку і заданням першої операції в новій базі `restore_values`. Передавати від однієї бази до іншої можна тільки значення параметрів. Якщо параметром операції `chain` буде порожній рядок, то буде виконано перезапуск поточної бази знань.

`do` - виконує передачу контролю до нового розділу, дозволяє виконувати керований пошук інформації між розділами. Іншими словами це не що інше, як пряме виведення. Синтаксис :

```
<do> ::= do <section-name>
```

`do_section_of` - як аргумент нею використовується параметр типу `category`. Ця акція необхідна, як альтернативна до оголошення типу :

```
if <category parameter name> = <option1> do <option1>
```

```
if <category parameter name> = <option2> do <option2>
```

...

Замість великого списку вибору можна записати просто :

```
do_section_of <category parameter name>
```

В процесі консультації саме значення певного параметра визначає, до якого розділу буде передано керування. Тому система намагатиметься визначити значення цього параметра першочергово.

`exit` - ця дія перериває процес консультації по поточній базі знань. Синтаксис :

```
<exit> ::= exit
```

`stop` – необхідна для оптимізації використання правил в розділі. Виконання цієї дії свідчитиме про те, що в розділі виконуватимуться жодні дії і не обчислюватимуться жодні логічні вирази. Дія `stop` допомагає інженеру зі знань уникнути повторного виконання складних умов. Синтаксис :

```
<stop> ::= stop
```

Параметри виконують роль змінних, які контролюють передачу управління між розділами. Будь-який параметр складається із поля оголошення і поля типу. Для опису окремих типів параметрів можна використовувати додаткові поля. Параметр може бути одним із таких чотирьох типів:

- `boolean parameter` (логічний параметр);
- `text parameter` (текстовий параметр);
- `number parameter` (числовий параметр);

- `category parameter` (параметр категорії).

Будь-який параметр може отримати своє значення одним із таких способів :

- в результаті відповіді на запитання системи ;

- як результат кількох правил;

- як результат виконання певних дій (actions).

Логічні параметри використовуються, коли параметр обмежується однією із величин - істина, хиба, або невизначено, частота появи, степінь підтвердження. За замовчуванням система автоматично генерує список вибору (listbox), що відповідає даним величинам.

З допомогою параметрів `declaration field` задають ім'я параметра і текст опису. Текст опису використовується системою для побудови відповідей на запитання користувача типу HOW(ЯК) and WHY(ЧОМУ).

Висновки

Деталізовано і узагальнено застосування метричних представлень, що компонуються і нечітких тверджень для формалізації множини висловлювань при надходженні інформації від гетерогенних кейс-базованих джерел із накладеними обмеженнями, що описують вуглеводневі поклади. На основі аналізу метричних ознак виділення вуглеводневих покладів запропоновано методику поетапної формалізації знань експерта з використанням коефіцієнтів впевненості в значеннях істинності висловлювань та тверджень. Представлено рішення з оптимізації обробки входжень в базі кейсів інтелектуальної системи за наявності слабкоструктурованої і неструктурованої інформації за рахунок процедур фільтрації за накладеними множинами обмежень і представлення даних з допомогою розподілу можливостей в заданій ієрархії обмежень. Побудовано гібридні структури знань та даних, що дозволило подолати розрив між представленням знань в кейс-базованій системі і процедурами інференції та солвінгу на множині даних. Розширено можливості процедур інференції по верифікації входжень істинності та хибності на основі гібридного підходу з використанням кейс-базованих запитів при наявності слабкоструктурованості та нечіткості входжень даних предметної області. Узагальнено та інстанційовано використання абстрактних потоків для обробки предикатних конструкцій та абстрагованих типів даних. На базі стандартних предикатних конструкцій побудовано механізм інференції для інтелектуальної системи. Розроблено методику видобування і обробки даних та знань в ході ініціалізуючого діалогу «експерт – база кейсів - інтелектуальна система».

Література

- 1 Гаврилова Т.А. Как стать инженером по знаниях / Т.А. Гаврилова // ВКНИИ, 21-25 ноября 1988г., Переславль-Залесский. - М., 1988. - С.332-337.
- 2 Загоровский И. М. Взаимодействие методов прямого приобретения знаний и обучения на примерах / И. М. Загоровский, В. И. Шалак // КИИ-92. - Тверь, 1992. - С.104-105.
- 3 Минский М. Структура для представления знания / М. Минский // Психология машинного зрения / Под ред. П.Уинстона. - М.: Мир, 1987. - С.249-338
- 4 Вудс У. А. Основные проблемы представления знаний / У. А. Вудс // ТИИЭР. - Т.74, №10. - С.32-46.
- 5 Городецкий Б. Ю. Зевахина Извлечение знаний из текста. Тезаурус текста / Б. Ю. Городецкий, Т. С. Зевахина и др. // КИИ-92. - Тверь, 1992. - С.101-102.
- 6 Quinlan J. Inferno: a cautious approach to uncertain inference // Comput.J. - 1983. - Vol.26. - P.255-269.
- 7 Шер А. П. Согласование нечетких экспертных оценок и функция принадлежности в методе размытых множеств / А. П. Шер. - Владивосток, ДВНЦ АН СССР, 1978. - 143с.
- 8 Инфанов А. П. Формирование концептуальной модели предметной области / А. П. Инфанов, Г. С. Осипов // Вопросы кибернетики. Проблемно-ориентированные вычислительные системы. - М., 1987. - С.103-111.
- 9 Ларичев О. И., Выявление экспертных знаний / О. И. Ларичев, А.И. Мечитов, Е. И. Мошкович, Е.М. Фуремс. - М.: Наука, 1989. - 128 с.
- 10 Zimmermann H.J. Fuzzy set, Decision-Making and Expert System Dordrecht: Kluwer Academic Publ., 1987. - 352 p.
- 11 Финн В. К. Правдоподобные выводы и правдоподобные рассуждения / В. К. Финн // Итоги науки и техники. Сер. теория вероятностей, математическая статистика, теоретическая кибернетика. Т.28. - М.: ВИНТИ, 1988. - С.3-84.
- 12 Кафаров В.В. Принцип описания химико-технологических процессов. Системный анализ процессов химической технологии. Применение метода нечетких множеств / В.В. Кафаров, И. Н. Дорохов, Е.П. Марков. - М.: Наука, 1986. - 390 с.
- 13 Аверкин А. Н., Нечеткие множества в моделях управления и искусственного интеллекта / А. Н. Аверкин, И. З. Батыршин [и др.] - М.: Наука, 1986. - С.139-169
- 14 Комаров С. И. Построение моделей предметных областей. Ч.2. Прямое приобретение знаний в системе SIMER / С. И. Комаров, Е. П. Куршев, Г. С. Осипов // Изв. АН СССР. Техн. кибернетика. - 1991. - №3. - С.192-197.
- 15 Юрчишин В. М. Проблеми впровадження експертних систем в нафтогазовидобувній промисловості / В. М. Юрчишин, В. І. Шекета // Доповідь на міжнародному симпозиумі «Комп'ютери в Європі: минуле, сучасне, майбутнє». - Київ. - 1998. - С. 367-372.
- 16 Юрчишин В. М. Прогнозування нафтогазових колекторів з використанням персонального комп'ютера / В. М. Юрчишин, В. І. Шекета // Тези доповіді на науково-практичній конференції. - Полтава. - 1998. - С.127-128.
- 17 Юрчишин В. М. Розробка моделі інформаційних потоків нафтогазовидобувного підприємства / В. М. Юрчишин, Г. О. Жученко, М. В. Крихівський, В. І. Шекета // Тези доповідей на науково-технічній конференції професорсько-викладацького складу університету. - Част.3. - Івано-Франківськ. - 1996. - С. 133
- 18 Шекета В.І. Особливості розробки інтерфейсу експертної системи для прогнозування нафтогазових колекторів / В. І. Шекета // Тези доповідей на науково-технічній конференції професорсько-викладацького складу університету. - Част. 2, Івано-Франківськ. - 1997. - С.190
- 19 Шекета В. І. Формування баз знань для прогнозування нафтогазових колекторів / В. І. Шекета, В. М. Юрчишин // Розвідка і розробка нафтових і газових родовищ. Серія: Технічна кібернетика та електрифікація об'єктів паливно-енергетичного комплексу. - Вип. 33. - Івано-Франківськ. - 1996. - С. 52-55.
- 20 Юрчишин В. М. Перспективи застосування інформаційно-математичного моделювання в нафтогазовидобувній промисловості / В. М. Юрчишин, М. В. Крихівський, В. І. Шекета // Математическое моделирование: Сб. науч. тр. / НАН Украины. Ин-т. математики. - Киев. - 1996. - С. 254-257.
- 21 Юрчишин В. М. Методика визначення коефіцієнта впевненості при виборі правил в базі знань / В. М. Юрчишин, В. І. Шекета, Ю. П. Кропельницький // Математические моделирования и информационные технологии. Сб. научн. трудов НАН Украины. Ин-т математики. - Киев. - 1998. - С.275-276.
- 22 Мальцев А. И. Алгебраические системы / А. И. Мальцев. - М.: Наука, 1988. - 234с.
- 23 Юрчишин В. М. Інформаційне моделювання нафтогазових об'єктів: монографія / Юрчишин В. М., Шекета В. І., Юрчишин О. В. - Івано-Франківськ: Івано-Франківський національний технічний університет нафти і газу, 2010. - 196 с.
- 24 Вовк Р.Б. Прийняття рішень у нафтогазовій галузі інструментами теорії обмежень: монографія / Р.Б. Вовк, О.Ф. Козак, В.І. Шекета, В.М. Юрчишин. - Івано-Франківськ: ІФНТУНГ, 2017 - 208 с.

Стаття надійшла до редакційної колегії
14.09.18

Рекомендована до друку
професором **Юрчишином В.М.**
(ІФНТУНГ, м. Івано-Франківськ)
д-ром техн. наук **Кузем М.В.**
(Івано-Франківський інститут права
імені Короля Данила Галицького,
м. Івано-Франківськ)