

Таким чином, знання використовуються як базові елементи вирішення завдань в ПО, а їх комбінація описує шлях вирішення задачі. При цьому композиція знань дає не тільки формальну схему – як композиція екстенціоналів, але й смислову – на рівні використання послідовності інтенціоналів. В свою чергу, коннотація дозволяє додатково використати коннотаційні зв'язки для того, щоб при необхідності – існування невизначеності – для вирішення окремого завдання використати поняття та знання, які відносяться до пов'язаних об'єктів ПО.

## ПЕРЕЛІК ЛІТЕРАТУРИ

1. Мейтус В.Ю. Введение в теорию интеллектуальных систем. Основные представления. Palmarium academic publishing, 2015. -189 с. ISBN 978-3-659-60297-9
2. Мейтус В.Ю. Проблемы разработки интеллектуальных систем // 21-я Международная Крымская конференция “СВЧ-техника и телекоммуникационные технологии”, 12-16 сентября 2011 г. Том 1. - 2011. - С.43-46. ISBN 978-966-335-351-7
3. Логический подход к искусственному интеллекту: от модальной логики к логике баз данных. Пер.с франц. / Тейз А., Грибомон П., Юлен Г. и др. - М.: Мир, 1998. - 494 с. ISBN 5-03-002519-7

УДК 004.896

**МЕЙТУС В.Ю., ЦЕПКОВА Д.П.**  
МННЦ ІТС НАНУ та МОНУ (Україна)  
НУ «Києво-Могилянська академія» (Україна)

### БАЗИ ЗНАНЬ З НЕПОВНОЮ ІНФОРМАЦІЄЮ ДЛЯ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

*Доповідь присвячена визначенню та обговоренню проблем, які пов'язані з неповною інформацією, що зберігається в базі знань інтелектуальної системи. Знання розглядаються як засіб, за допомогою якого інтелектуальна система записує, зберігає та відтворює опис оточуючого середовища (предметної області). Для визначення знань використовується чотирьохкомпонентна модель подання знання, яка для кожного описаного елемента предметної області зберігає його характеристики та смисл. Але інформація про елемент може бути неповною. Можуть бути відсутні окремі компоненти чи не подані функціональні залежності, або використовуватися різні структурні моделі в предметній області при побудові знань. Тому в базу знань включаються додаткові функції узгодження, поповнення та приведення знань.*

Загальна схема побудови інтелектуальної системи (ІС), що подана в роботі [1], ґрунтується на припущенні, що діяльність ІС завжди визначається предметною областю (ПО), в якій система повинна вирішувати послідовність завдань, які визначають поведінку системи в цій області. Інтелект тоді розглядається як суттєва властивість, що забезпечує ефективну можливість пошуку і побудови рішень для цих завдань.

Кожна ПО в навколишньому світі нескінченно різноманітна за своїми властивостями і якість. Взаємодія з нею заснована на можливості подання цієї області у вигляді деякої моделі. Природно, що самі моделі можуть відрізнятися між собою, і кожна відображати тільки деякі риси ПО, проте достатні для вирішення окремих класів задач. Для ІС модель замінює ПО, а зв'язок між моделлю і ПО задається інтерпретацією моделі в ПО. В основі моделі лежить інформація, що сприймається з області інтелектуальною системою і використовується нею при побудові моделі. Ця інформація надається у вигляді знань, які розглядаються як складові елементи при створенні моделі ПО.

Будь-яке завдання в ПО подається як пара точок (або областей) в просторі, пов'язаному з моделлю і інтерпретованому в ПО. Одна з цих точок являє собою умову задачі, а друга - результат її вирішення, хоча іноді ці області співпадають одна з одною. Саме ж рішення розглядається як шлях в ПО, що зв'язує ці області – умову задачі з її рішенням. Зазвичай це рішення шукається спочатку в моделі ПО, а потім переноситься (розглядається) як рішення задачі в ПО. Якщо модель була адекватна ПО, то і рішення, отримане за моделлю, відповідає цій області і виявляється рішенням завдання, що відноситься до ПО.

*Як знання будемо розглядати сукупність фактів, форм, тверджень (одне або одночасно декілька), уявлень, схем, моделей, які задають об'єкти ПО, образи, структури і комбінації елементів в ПО, структурні зв'язки (залежності, відносини) між об'єктами і структурами, геометрію ПО. Ще раз підкреслимо, що знання - це описи окремих складових ПО, а сукупність*

знань про ПО - це уявлення цієї області, на підставі якого будується її модель і формується поведінка ІС.

В роботі [1] пропонується розглядати знання у вигляді іменованої чотирьохкомпонентної структури, що включає для кожного знання, що зберігається в БЗ, основні компоненти подання знань: *тип, інтенціонал, екстенціонал і конотацію* знання. Для завдання типів застосовується алгебра типів, а для кожної з решти компонент використовується своя логічна мова. В основу цих мов покладений мову висловлювань  $L_0$ , в якій висловлювання записуються у вигляді логічних формул [2].

Оскільки знання відіграють основну роль при побудові моделі ПО і знаходженні вирішення завдань, тобто при виконанні основних функцій ІС, то знання необхідно отримувати, зберігати, коригувати та видавати при формуванні поведінки системи. Для цього в ІС існує база знань (БЗ), яка складається з двох компонент - банку знань (БкЗ) і системи управління базою знань (СУБЗ). У банку знань зберігаються самі знання, а в СУБЗ виконуються всі операції, пов'язані з управлінням всією сукупністю знань. Залежно від того, як представлені знання, визначаються і операції над ними, як і сам процес використання та композиції знань при вирішенні задач інтелектуальною системою.

У той же час при використанні знань виникає ряд проблем, що спостерігаються, наприклад, при аналізі когнітивної діяльності людини. Однією з таких проблем є неповнота знань. Дослідження показують, що ця неповнота знань є характерною властивістю всієї сукупності знань, що використовуються людиною при формуванні своєї поведінки. Найчастіше вона пов'язана з тим, що інтерпретація, яка визначає семантику знань, може задаватися різними структурними моделями, які люди використовують як для спілкування, так і для розуміння інформації, поданої в знаннях. А відмінності в моделях, використаних при інтерпретації, призводять до різниці смислових характеристик однакових по імені (найменуванню) понять. Цим можна пояснити, наприклад, існуючі відмінності у визначеннях поняття «інтелекту», які зустрічаються в літературі [1].

Крім того, неповнота знань пов'язана з тим, що окремі зв'язки між подібними поняттями можуть бути відсутніми, або частково не збігатися при вживанні тих чи інших знань в певному контексті. Тоді при використанні знань в різних процесах або при вирішенні різних завдань виникають формальні відмінності таких процесів, які, в свою чергу, можуть призводити до різних варіантів рішення задачі, до різних наслідків, що випливають з такого рішення. І якщо в математиці в ході формального доказу такі відхилення зазвичай мінімальні, хоча і зустрічаються, то в інших науках, особливо соціальних, дослідники часто приходять до зовсім відмінних висновків.

Нарешті, неповнота знань в базах знань при чотирьохкомпонентному поданні знання може виражатися у відсутності окремих компонент для деяких знань. Наприклад, заданий екстенціонал для окремого знання, тоді як інтенціонал - смисл знання - відсутній. Або смислова складова знання в повному обсязі не визначена. Або існуюче в ПО відношення не показано в поданні знання, і цей розрив не дозволяє здійснити необхідний перехід при побудові рішення задачі. Подібні проблеми часто зустрічаються і в людських знаннях, що деякий час не дозволяє вирішувати окремі завдання, які потім успішно вирішуються дослідниками, які звернуть увагу на відсутність відповідних елементів.

В загальному випадку СУБЗ на відмінність від СУБД крім операцій, пов'язаних з пошуком, поповненням, модифікацією знань, тобто операціями з множиною знань, поповнюються операціями, які аналізують знання відносно всієї ПО, та корегують (модифікують) окремі знання при змінах інших знань. Виникають багатолістові знання, при яких окремі множини знань поєднуються з областями істинності, в яких істинними стають визначені екстенціонали.

У кожному можливому випадку неповноти знань СУБЗ поповнюється додатковими операціями, які дозволяють корегувати існуючі знання відповідно тих умов, які визначають неповноту бази знань. При цьому в СУБЗ може також використовуватися своя логіка виводу, наприклад, немонотонна логіка, а також додаватися різні аксіоми поповнення БЗ. Тому можна зробити висновок, що бази знань не тільки відрізняються від баз даних, але й включають спеціальні механізми прив'язки знань не тільки до ПО, але й до суб'єкта, з яким асоційована ІС, якого моделює ця система. В цих умовах неповнота інформації визначає необхідність використання додаткових механізмів побудови бази знань, які пов'язані не тільки предметною областю, але й принципами створення ІС.

## ПЕРЕЛІК ЛІТЕРАТУРИ

1. Мейтус В.Ю. Введение в теорию интеллектуальных систем. Основные представления. -Palmarium academic publishing, Саарбрюкен 2015. - 189 с. ISBN 978-3-659-60297-9
2. Логический подход к искусственному интеллекту: от модальной логики к логике баз данных. Пер.с франц. / Тейз А., Грибомон П., Юлен Г. и др. - М.: Мир, 1998. - 494 с. ISBN 5-03-002519-7

УДК 004.056.55

*МИРОНІВ І.В., ЖИХАРЕВИЧ В.В., ОСТАПОВ С.Е.*

Чернівецький національний університет імені Юрія Федьковича (Україна)

### АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ СИМВОЛІВ ТЕКСТУ

Розпізнавання символів різних зображень забезпечує рішення ряду наукових і прикладних задач при ідентифікації об'єктів різної природи. Сучасні методи розпізнавання символів використовуються для вирішення як типових завдань, наприклад розпізнавання тексту, так і спеціалізованих завдань, орієнтованих на розпізнавання символічної інформації, нанесеної на поверхню різних об'єктів. В даний час існує достатньо велика кількість програм, призначених для розпізнавання тексту (наприклад, FineReader, Readiris, ScanSoft OmniPage, CuneiForm та ін.). Кожна з цих програм пропонує свою реалізацію рішення задачі обробки та розпізнавання зображень. В основному ці програми є комерційними, тому методи вирішення завдань, закладені в них, відомі тільки їх розробникам, і практично неможливо визначити для яких завдань вони підходять і які завдання їм не під силу. Крім того, всі ці програми поставляються у вигляді виконуваних модулів, що робить неможливими аналіз працездатності програм, якості їх роботи і модифікацію використаних математичних моделей і алгоритмів.

У цій статті розглянемо найбільш відомі і поширені методи розв'язання задачі розпізнавання символів. На даний момент виділяють три основні підходи для вирішення сформульованого завдання: структурний, ознаковий і шаблонний. Кожному методу притаманні свої достоїнства і недоліки. Розглянемо докладно кожний метод окремо.

#### **Розпізнавання по шаблонах**

Програмне забезпечення *OCR (Optical Character Recognition* - оптичне розпізнавання символів) зазвичай працює з великим растровим зображенням сторінки з сканера. При цьому більшість систем має шаблони, створені для різних накреслень. Після декількох розпізнаних слів, програмне забезпечення визначає використовуваний шрифт і шукає відповідні пари тільки для цього шрифту. В деяких випадках програмне забезпечення використовує чисельні значення частин символу (пропорцій), щоб визначити новий шрифт. Це може поліпшити ефективність розпізнавання.

#### **Структурний підхід**

Система розпізнавання OCR - Care Omni Page Professional використовує алгоритм, заснований на знаходженні загальних специфічних особливостей символів.

Ця система містить 100 різних алгоритмів для ідентифікації 100 різних символів: верхнього і нижнього регістра від «А» до «Z», записи чисел і символів пунктуації. Кожен з цих алгоритмів шукає «Особливості» накреслень типу «острів», «півострів», точок, прямих відбитків і дуг. Експертні системи також розглядають горизонтальні і вертикальні проекції відбитки букви і звертають увагу на основні особливості у створених кривих, підсумовуючи в них число темних пікселів.

#### **Ознакові методи**

Ознакові методи базуються на тому, що зображенню ставиться у відповідність  $N$  –мірний вектор ознак. Розпізнавання полягає в порівнянні його з набором еталонних векторів тієї ж розмірності. Завдання розпізнавання, прийняття рішення про приналежність образу того чи іншого класу, на підставі аналізу обчислених ознак, має цілий ряд строгих математичних рішень в рамках детерміністичного і імовірного підходів [3].

У системах розпізнавання символів найчастіше використовується класифікація, заснована на підрахунку евклідової відстані між вектором ознак розпізнаваного символу і векторами ознак еталонного опису. Тип і кількість ознак в більшій мірі визначають якість розпізнавання. Формування вектора виробляється під час аналізу попередньо підготовленого зображення. Даний