

УДК 004:336

**М. ШВЕЦЬ**, член-кореспондент АПрН України,**Ю. КЛІМАШЕВСЬКА**, завідувач лабораторією НДЦПІ АПрН України

## ОСНОВНІ ЗАСАДИ ПОБУДОВИ БАЗ ЗНАНЬ У ПРАВОВІЙ СФЕРІ

*Анотація.* В статті розглядаються проблеми формалізації знань у правовій сфері, описуються можливості використання базових моделей подання знань з метою побудови систем підтримки прийняття рішень у цій сфері.

Довгий час генеральним напрямом розвитку ЕОМ вважались проекти, спрямовані на розроблення машин з максимальною швидкістю та гранично великим обсягом пам'яті. Удосконалення елементної бази хоч і привело до величезного скачка у швидкості ЕОМ та використовуваної нею пам'яті, очікуваного зростання “інтелекту” комп'ютера не сталося. Комп'ютери продовжували залишатися “невігласами”. Крім того ЕОМ продовжували залишатися у великій ізоляції від переважної частини людства, через що зростала роль людей – програмістів, які перекладали на зрозумілу для ЕОМ мову все те, що хотіли б одержати від ЕОМ інші спеціалісти. Виникла необхідність зруйнувати такий бар'єр між ЕОМ та користувачем. Наука поставила перед собою завдання: зробити так, щоб кожна людина – спеціаліст могла б ставити перед ЕОМ завдання та одержувати їх вирішення зрозумілою для себе мовою.

Подальший розвиток комп'ютерної техніки відбувався (та й відбувається надалі) саме у напрямку створення суперкомп'ютерів, які забезпечували б найбільш зручний і природний спосіб взаємодії користувачів з ЕОМ. Суттєво підвищилась ефективність ЕОМ за рахунок розвинутих програмних засобів, використання мов програмування високого рівня, наближених до природної мови, удосконалення операційних систем. Але слід зазначити, що роботи в цьому напрямку суттєво стримуються проблемою забезпечення змістовного аналізу вхідних даних. Через те, в основному, послуги, що надають електронно-обчислювальні машини, носять технічний характер, в той час як творча, наукова, аналітична діяльність здійснюється людиною. Так виникла ідея створення ЕОМ п'ятого покоління. Перший проект такої машини подала Японія [1]. Слід зазначити, що до того часу спеціалістами європейських країн, США, Канади, СРСР (в тому числі і України) уже були зроблені серйозні напрацювання в галузі штучного інтелекту. І хоч мета, задекларована в японському проекті, не досягнута й по сьогодні, всі розвинені країни найсерйознішим чином продовжують аналогічні дослідження та проекти у себе. Зростання потужності обчислювальної техніки знову й знову відроджує сподівання швидкого та ефективного моделювання інтелектуальних можливостей людини. Але багаторічний досвід використання обчислювальної техніки переконує, що ефективні моделі можна побудувати тільки там, де є чіткий та однозначний інформаційно-логічний аналіз предметної області.

Багато чого з того, чим володіє людина в якості знань, досвіду й інформації, є результатом інтуїції. Тривала спеціалізація у тому чи іншому виді діяльності дозволяє розширити наші інтелектуальні здібності у відповідній предметній області. При цьому ми користуємося великим запасом неусвідомлених знань, навиків і вміння, що сформувалися на протязі тривалої еволюції людства в конкретних умовах існування. Розроблення комп'ютерів, здатних моделювати аналітичне мислення людини, вимагає принципово нових способів

організації й обробки інформації. Найважливішими шляхами зростання рівня “інтелектуального розвитку” комп’ютерів а також способів їх взаємодії з людиною є розроблення засобів розуміння машиною природної мови та створення баз знань.

Термін “знання” для систем штучного інтелекту настільки ж звичний, як і термін “дані” для перших поколінь електронно-обчислювальних машин. На перших порах використання ЕОМ дані виконували роль “їжі для вічно голодних програм”. Але з часом їх роль невпинно росла, їх структура ускладнювалась. Від машинного слова, що містилося в одній комірці пам’яті ЕОМ, відбувався перехід до векторів, масивів, файлів, списків. Нарешті, виникли абстрактні типи даних, за допомогою яких стало можливим створювати такі структури даних, які дозволили найефективніше вирішувати задачі в тій чи іншій предметній галузі. Огляд всього шляху розвитку математичного програмного забезпечення ЕОМ свідчить про те, що між даними з розвиненою структурою та знаннями нема абсолютно чіткої межі.

Подамо визначення (наприклад, згідно з [2]) деяких базових понять у галузі штучного інтелекту та теорії побудови експертних систем.

*Предметна галузь* – це деяка частина реального, уявного чи видуманого світу, що є предметом зацікавленості людини-дослідника. Іншими словами – це всі ті сутності у сфері інтересів людини-дослідника, які існували, існують чи могли б існувати. Під *сутностями* мається на увазі все, що завгодно, якщо це “що завгодно” є предметом зацікавленості людини і сприймається нею як дещо суттєве (і таке, що існує, або може існувати) незалежно від того, існує воно матеріально, в абстрактній формі чи тільки в уяві. Процес дослідження, вивчення предметної галузі, формування її уявного образу, що завершується формальним поданням цього уявного образу у вигляді бази знань, називається концептуальним моделюванням. Отже, якщо трактувати *знання (уявний образ предметної галузі)* як відображення (краще чи гірше, більше чи менше адекватне) об’єктивної реальності у свідомості людини, то *база знань* – це формальне подання (з використанням засобів формальної мови) цього уявного образу в рамках інформаційної системи.

Згідно з роботами [3 – 6] знання класифікуються наступним чином:

- предметні знання – якісні характеристики та показники об’єктів та елементів, з яких складаються ці об’єкти;
- концептуальні знання – множини понять, описів об’єктів, їх властивостей та взаємозв’язків, що використовуються в даній предметній області;
- конструктивні знання – знання про взаємодію об’єктів проблемної області та їх структуру;
- процедурні знання – способи та методи виконання дій над об’єктами та їх описами.

Крім того знання характеризуються такими властивостями: інтерпретованістю, структурованістю, зв’язністю, активністю.

Але, якщо уважно проаналізувати ті структури, які називають даними, то можна побачити, що вони теж мають подібні властивості. Сучасні системи керування базами даних (СКБД) також забезпечують інтерпретацію всіх інформаційних одиниць (даних), що зберігаються в БД. Розглянемо приклад зберігання в пам’яті ЕОМ даних, що містяться в обліковій картці відділу кадрів.

Номер справи	26	27	28
Прізвище	Іванченко	Петренко	Миколаєнко
Спеціальність	слюсар	токарь	токарь
Вік	28	35	25

Якщо те, що знаходиться в останніх трьох стовпчиках, міститься в пам'яті машини у вигляді машинних слів, то код числа “28” не залежить від того, яке змістовне навантаження воно несе, і є зовсім не інтерпретований. Тільки програміст знає, за якою адресою в пам'яті машини міститься номер справи, а за якою – вік Іванченка. Якщо ж в пам'яті зберігається весь масив даних з таблиці, то на запити до ЕОМ: “скільки співробітників, вік яких 28 років?”, та “скільки років співробітнику, номер справи якого 28?” користувач обчислювальної системи одержить різні, змістовно правильні відповіді.

Як знання, так і дані в сучасних системах керування базами даних характеризуються структурованістю. Знання діляться на описи об'єктів, процесів, ситуацій, явищ, а також на опис способів їх подання та модифікації. Так само й серед типів даних, що допускаються сучасними засобами програмування, є такі типи як СТРУКТУРА, КЛАС і т. ін.

Дещо суттєвіша відмінність між даними та знаннями полягає в такій характеристиці як зв'язність. Певною мірою зв'язки типу “рід – вид”, “елемент – клас”, “тип – підтип” властиві і знанням і сучасним типам даних. Разом з тим сучасні бази знань оперують такими відношеннями між окремими елементами як “одночасно”, “бути в тому ж місці простору”, “бути в певній ситуації” і т. д.

Рисою, що суттєво відрізняє знання від даних, є їх активність. Активність знань передбачає цілеспрямоване використання інформації. У людини знання відіграють активну роль. Їх взаємодія може породжувати нові знання, спонукати до здійснення тих чи інших дій, спричиняти виконання певних процедур.

Таким чином, знання та дані мають багато спільного. Але знання мають більш складну структуру. Їх іноді називають мета даними, або даними про дані.

Бази знань створюються і розвиваються в рамках експертних систем (ЕС), основна мета створення яких – накопичення, систематизація та формалізація знань у даній предметній галузі та подання їх спеціалістам для здійснення наукової та практичної діяльності. В базах знань в комп'ютеризованій формі подаються знання людини у певній конкретній предметній галузі.

При проектуванні та розробленні експертних систем значну увагу слід приділяти створенню інструментарію баз знань та експертних систем згідно з класами застосування:

- засоби управління базами знань
- методи та засоби подання знань
- засоби діалогу.

Вирішення поставлених задач вимагає від користувачів ЕС доброї підготовки не тільки в предметній галузі, але й у галузі інформатики та математики. Тому однією з основних проблем при створенні сучасних ЕС є виділення такої підмножини природної мови, яка б характеризувалася точністю й однозначністю формальних мов і при цьому була б достатньо простою та зрозумілою в користуванні.

Узагальнення, систематизація та формалізація знань є настільки важливим, наскільки й важким завданням в галузі створення експертних систем, у зв'язку з цим на сучасному етапі інформатизації суспільства велике значення має проблема накопичення та використання знань. При цьому особлива роль належить спеціалістам, які володіють знаннями в різних предметних галузях і які здатні свої знання спрямовувати не тільки на вирішення конкретних проблем даної предметної галузі, а й на формалізацію предметних знань (понять та відношень між ними), а також на формалізацію правил одержання нових знань.

На розроблення тих чи інших моделей подання знань значний вплив мають психологічні дослідження, спрямовані на виявлення тих структур, у вигляді яких людина

зберігає інформацію про оточуючий її світ у своїй свідомості та про організацію своєї діяльності у цьому світі. Система подання знань складається із наступних компонент: мови подання знань, механізму набуття знань, саме знань, алгоритму роботи зі знаннями та механізму одержання виводів. Таким чином, під *системою подання знань* слід розуміти засоби, які дають можливість описувати знання про предметну область за допомогою формальної мови, організовувати накопичення знань, їх аналіз, узагальнення, структурування та зберігання, а також одержувати нові знання з наявних, оперувати з неповними та неточними знаннями, здійснювати пошук потрібної інформації та перевірку знань на не суперечливість.

Подання знань – одне з найбільш сформованих напрямів штучного інтелекту, до якого традиційно відносяться питання розроблення формальних мов і належних програмних засобів.

В сучасних інтелектуальних системах використовуються наступні моделі подання знань:

- логічні
- продукційні
- семантичні мережі
- фреймові
- об’єктно орієнтовані
- у відповідності з біонічним напрямом розвитку інтелектуальних систем знання подаються на нейроподібних зростаючих мережах.

**Логічні схеми подання знань** використовують поняття *константи, змінної, функції, предиката, логічної функції і квантора* для представлення фактів у вигляді логічних формул за деякою логікою (першого або більш високого порядку, багатозначної, модальної, нечіткої і т. п.). При такому підході база знань являє собою сукупність логічних формул, що забезпечують частковий опис стану. Модифікації бази знань відбуваються при додаванні або видаленні логічних формул.

Для того, щоб можна було використати методи логіки предикатів для конкретної предметної галузі, спочатку необхідно докладно проаналізувати структуру цієї галузі. В результаті аналізу цього виділяється множина значущих сутностей. Ця множина називається областю інтерпретації. Якщо, наприклад, галузь є арифметика, то областю інтерпретації може бути множина натуральних чисел.

На наступному етапі визначаються важливі функції над елементами області інтерпретації, а також значущі відношення, які існують між її елементами.

Розгорнуту характеристику логічних моделей подання знань і одержання виводів зроблено в роботі [7]. Розглянемо приклад застосування логічних формул для відображення фактів, що описують спорідненість між деяким колом осіб. Введемо такі відношення та позначення:

БАТЬКО (x,y); x є батьком y –  $P_1(x,y)$ ;

МАТИ (x,y); x є матір’ю y –  $P_2(x,y)$ ;

БРАТ (x,y); x є братом y –  $Q_1(x,y)$ ;

БРАТ\_ПО\_БАТЬКОВІ (x,y); x є братом по батькові y –  $Q_2(x,y)$ .

Можна записати аксіоми, які описують спорідненість між людьми:

$\forall x_1 \forall x_2 \forall y \forall z (P_1(x_1, y) \wedge P_1(x_2, y) \wedge P_2(x_1, z) \wedge P_2(x_2, z)) \rightarrow Q_1(x_1, x_2)$

{якщо y є батьком  $x_1$  та  $x_2$ , а також z є матір’ю  $x_1$  та  $x_2$ , то  $x_1$  та  $x_2$  є братами}.

$$\forall x_1 \forall x_2 \forall y \forall z (P_1(x_1, y) \wedge P_1(x_2, y) \wedge ( ( P_2(x_1, z) \wedge \neg P_2(x_2, z) ) \vee ( \neg P_2(x_1, z) \wedge P_2(x_2, z) ) ) ) \rightarrow Q_2(x_1, x_2)$$

{якщо  $y$  є батьком  $x_1$  та  $x_2$ , і  $z$  є матір'ю  $x_1$  і не є матір'ю  $x_2$ , або  $z$  не є матір'ю  $x_1$  але є матір'ю  $x_2$  то  $x_1$  та  $x_2$  є братами по батьковій лінії}.

Нехай база знань містить такі відомості про осіб:

{ Андрій;  
Дмитро – батько Андрія;  
Марина – мати Андрія}.  
{ Олег;  
Дмитро – батько Олега;  
Ірина – мати Олега}.

На основі записаних вище аксіом можна формулювати запит і одержувати правильну відповідь стосовно спорідненості Андрія та Олега. Вони є братами по батьковій лінії – знання, яке явно не фігурує в базі знань.

**Продукційні моделі.** Правила продукцій (продукційні правила) в загальному випадку записуються так:

**ЯКЩО**  $P_1, P_2, \dots, P_n$  **ТО**  $B$ ,

де:  $P_1, P_2, \dots, P_n$  – умови застосовності продукції,  $B$  – висновок (дія).

Записане правило означає: “якщо всі умови від  $P_1$  до  $P_n$  істинні, то  $B$  – також істинне”, або “якщо всі умови від  $P_1$  до  $P_n$  істинні, то слід виконати дію  $B$ ”.

Наведемо приклад, як можуть бути використані продукційні правила для визначення покарання за злочин, передбачений статтею 146 Кримінального кодексу України.

**ЯКЩО ДІЯННЯ Є викрадення людини**

**КВАЛІФІКУЮЧА ОЗНАКА Є з корисливих мотивів**

**ТО ПОКАРАННЯ Є обмеження волі СТРОК Є до п'яти років**

**АБО ПОКАРАННЯ Є позбавлення волі СТРОК Є до п'яти років.**

**Семантичні мережі.** У найзагальнішій формі семантична мережа являє собою знання у вигляді сукупності об'єктів (вузлів) і бінарних сполучень (орієнтованих позначених ребер), перші з яких символізують індивідууми або визначені концепції, а другі – бінарні відношення між ними. Відповідно до цього база знань являє собою сукупність об'єктів і відношень, визначених над ними, а модифікація бази знань проводиться шляхом вставки або видалення об'єктів і їх відношень. Семантичні мережі здатні відображати структуру знань з усією складністю їх взаємозв'язків, пов'язувати в єдине ціле об'єкти і їх властивості. Важливою особливістю семантичних мереж є ієрархія. Вузол, який представляє певні знання, з'єднано дугами типу “Є” з вузлами вищого та нижчого рівнів (відношення виду та роду, відношення належності елемента до класу). Спосіб подання знань семантичною мережею дозволяє легко робити виводи завдяки ієрархії успадкування. Екземпляр нижчого рівня містить атрибути (властивості, характерні ознаки), які має екземпляр верхнього рівня (прототип). Саме ця властивість називається успадкуванням атрибутів між рівнями ієрархії “Є”. В семантичних мережах можна відображати також відношення “ціле – частина”.

На Рис. 1 наведено частину семантичної мережі, яка описує поняття “юридична особа приватного права”.

**Фреймові моделі.** Поняття фрейма вперше було запропоновано вченим Мінським [8] і відіграло ключову роль у дослідженнях у галузі інженерії знань. Фрейм – це спеціальні інформаційні структури для подання знань про об'єкти, які можна описати

деякою сукупністю понять і сутностей. Фрейм має певну внутрішню структуру, яка складається з елементів, що називаються слотами. Кожен слот, в свою чергу, може подаватися певною структурою даних, процедурою, або бути пов’язаним з іншими фреймами. Ця структура базується на відношеннях “абстрактне – конкретне” і характеризується тим, що на верхніх рівнях розміщені абстрактні об’єкти (концепти), а на нижніх – конкретні, причому об’єкти нижніх рівнів успадковують атрибути (властивості) об’єктів верхніх рівнів за умови, коли відповідні описи у них самих відсутні.

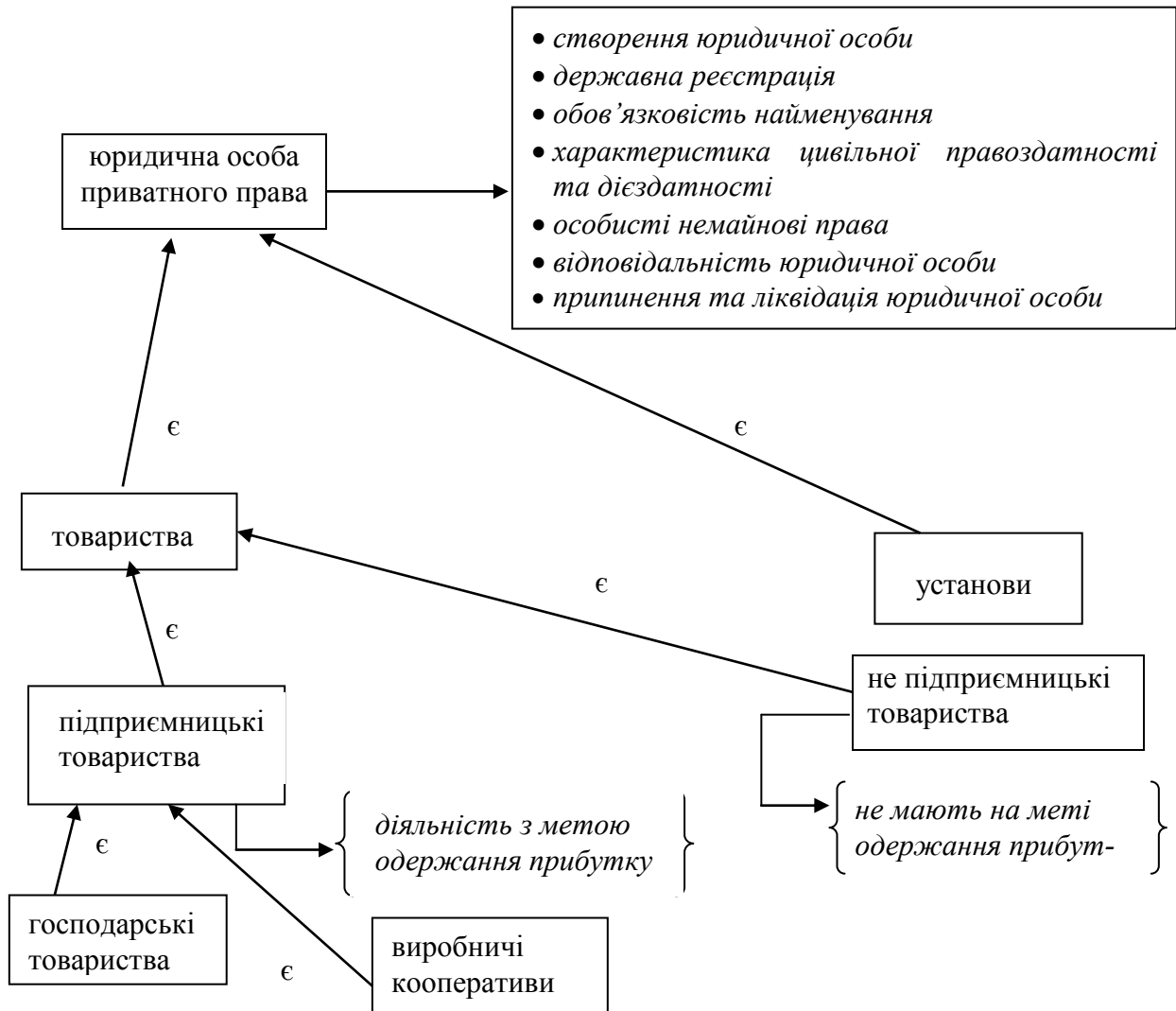


Рис.1

Для прикладу наведемо фрагмент схеми класифікаційної структури, пов’язаної з поняттям договору (див. Рис. 2). Наведений приклад показує схему успадкування (чи не успадкування) таких властивостей як “форма укладення договору” та “момент укладення договору”.

Існують також структури, які базуються на відношеннях “ціле – частина” і характерні тим, що об’єкти нижнього рівня є частиною об’єкта верхнього. У відношеннях “ціле-частина” атрибути не успадковуються. Наприклад, якщо об’єкт “кримінальна справа” складається з елементів “засуджений”, “суд, що виніс рішення у справі”, “дата винесення рішення” (та деяких інших), то елемент “суд, що виніс рішення у справі” не успадковує атрибути об’єкта “кримінальна справа”. Успадкування фреймами значення слотів

здійснюється в тому випадку, якщо у фреймі є слот **Різновид**, в якому міститься ім'я фрейма вищого рівня.

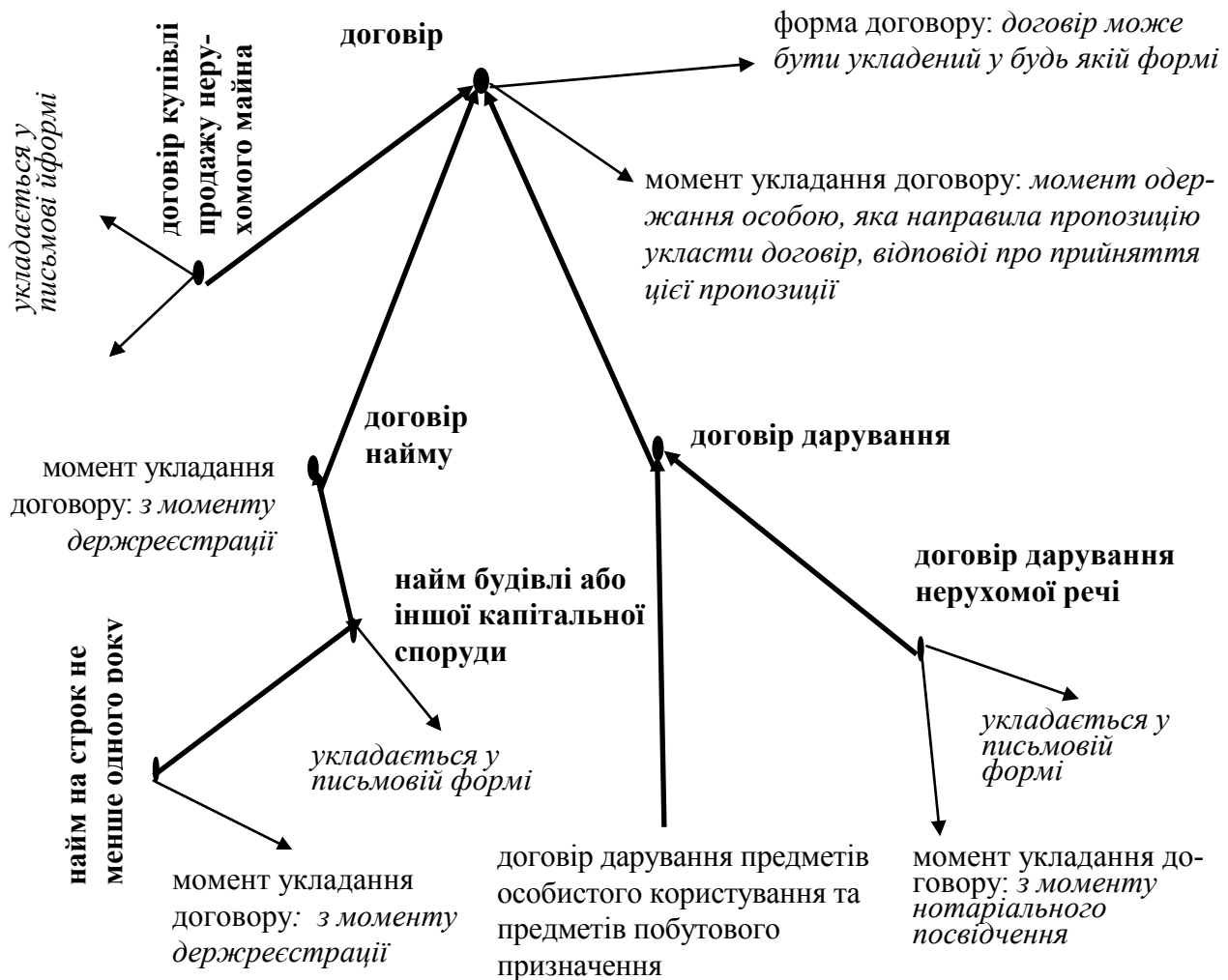


Рис. 2

Наведемо приклад фреймової моделі для подання відомостей про те, що у вівторок о 14.00 на Саксаганського 110 проводяться наукові семінари, а 14.11.2006 на цьому семінарі Іванишин і Петришин зроблять доповідь “Бази знань у галузі держави і права”.

**Фрейм:** НАУКОВИЙ\_СЕМІНАР

**Час :** Вівторок\_14.00

**Місце :** Саксаганського\_110.

**Фрейм:** БАЗИ\_ЗНАНЬ\_У\_ГАЛУЗІ\_ДЕРЖАВИ\_І\_ПРАВА

**Різновид:** НАУКОВИЙ\_СЕМІНАР

**Дата :** 14.11.2006

**Доповідачі :** Іванишин, Петришин.

В **об’єктно-орієнтованому** програмуванні основними складовими одиницями є об’єкти. Кожний об’єкт містить деяку структуру даних (або типів даних) і набір процедур (методів), які призначені для опрацювання цих даних. Об’єкти з однаковими властивостями і поведінкою об’єднуються в класи. Кожен об’єкт входить в один клас. Кожен клас має одного предка, який називається суперкласом (надкласом). Клас може мати од-

ного або декілька нащадків, які називаються підкласами. Є клас, який не має суперкласу і є коренем дерева ієрархії класів. Кожен клас успадковує змінні екземпляра і методи суперкласу. Крім того, він може містити нові змінні екземпляра і нові методи, може також перевизначити успадковані.

На Рис. 3 подано приклад ієрархії класів, яка несе інформацію про підприємства.

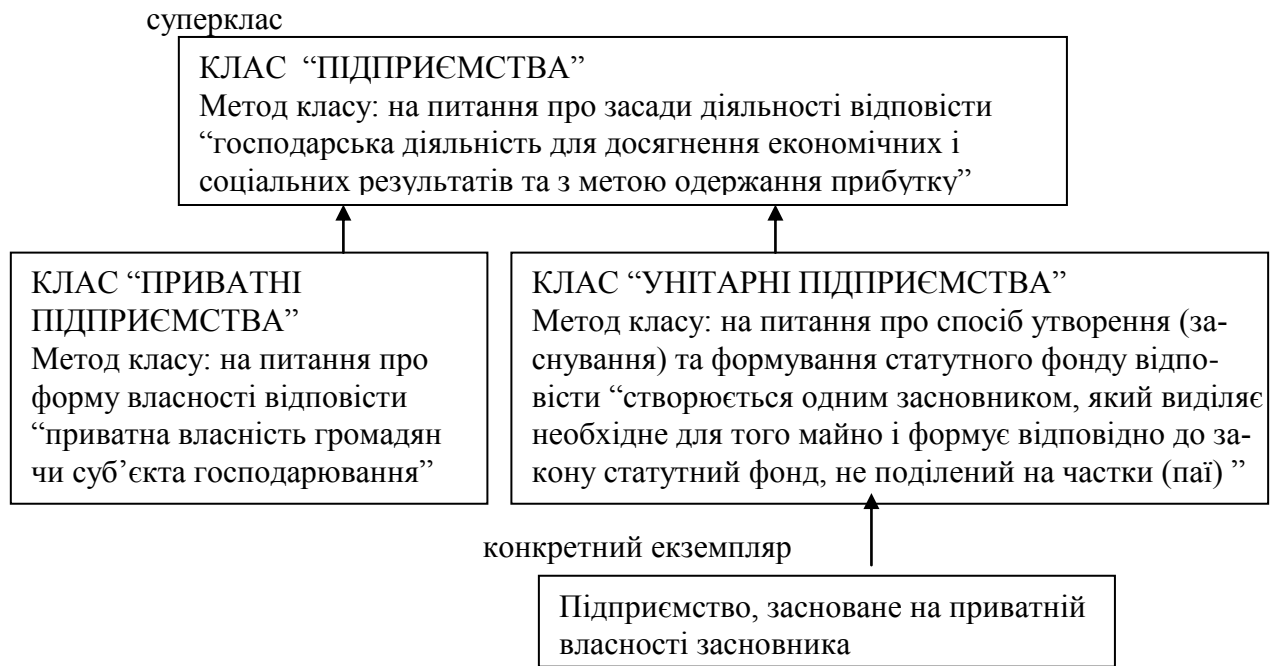


Рис. 3

Формалізм об'єктно-орієнтованого програмування можна трактувати як уточнення формалізму фреймів. У формалізмі фреймів не розрізняють види відношень клас/підклас і клас/конкретний екземпляр (тобто те саме відношення **Різновид** може існувати і як відношення між двома класами, і як відношення між класом і конкретним екземпляром). В об'єктно-орієнтованому програмуванні ці два види відношень розрізняються.

Зрозуміло, що від вибору моделі подання знань суттєво залежить ефективність розробок інтелектуальних систем. У зв'язку з цим було б добре забезпечити можливість роботи системи з різними моделями подання знань. Хоча слід зазначити, що об'єднання в одній системі блоків, що містять знання в різних поданнях, пов'язано з використанням складних процедур керування такими базами знань.

Кожний з формалізмів подання знань має свої переваги й недоліки. Вибір оптимального способу подання знань у кожному випадку значною мірою залежить від характеру та складності тих задач, які доводиться розв'язувати.

**Подання та використання нечітких знань.** Знання, з якими часто доводиться оперувати, не завжди можна точно описати. Часто бувають ситуації, коли стосовно того чи іншого судження людина може дати відповідь “можливо так, а можливо й ні; близько до істини у дев'яти випадках із десяти” і т. д. Крім того, під час вивчення предметної галузі та формального подання відповідного уявного образу людина може опинитися в ситуації, коли вона нічого не знає стосовно істинності того чи іншого судження. Такі знання називаються нечіткими, неповними.

Людям повсякденно доводиться розв'язувати різноманітні проблеми і робити виводи саме в середовищі нечітких знань. Для того, щоб у подібних ситуаціях можна було



використовувати інтелектуальні інформаційні системи і щоб ці системи вийшли за рамки одержання простих символічних висновків і наблизилися до мислення людини, необхідно мати відповідні методи подання нечітких знань і механізмів одержання висновків, придатних для такого середовища. Поки знання не формалізовані і певним чином не описані, вони не можуть бути використані в комп'ютерах. Нечіткості, з якими доводиться мати справу в інженерії знань, можна класифікувати [7] наступним чином:

- 1) не детермінованість висновків;
- 2) багатозначність;
- 3) ненадійність;
- 4) неповнота;
- 5) нечіткість чи неточність.

Але навіть у випадках нечітких знань людина – дослідник чи експерт – робить цілком визначені висновки. Методи, яким користуються в інженерії знань для одержання висновків при оперуванні нечіткими знаннями базуються на розумінні ймовірності деякої події, залученні до аналізу додаткової інформації. Одержані висновки подаються з відповідним значенням ймовірності.

Слід зазначити, що використання в інтелектуальних системах тих чи інших моделей подання знань є однією з основних і найскладніших тем, які відносяться до інженерії знань. Особливої уваги потребує проблема вивчення способів подання знань у правовій галузі. Побудова моделей у цій галузі пов'язана зі значними складностями, що викликані неможливістю визначити в межах одного окремо взятого методу моделювання всього розмаїття об'єктів, процесів та явищ, якими характеризується правова сфера діяльності людини. У зв'язку з цим дуже гостро стоїть завдання системного підходу до аналізу цієї предметної області з позицій правової інформатики. Враховуючи той факт, що значна кількість даних міститься в текстових документах, можливо, з метою сумісного використання необхідно узгодити як логіко-лінгвістичні так і формально-математичні методи обробки інформації, сприяючи тим самим розробленню систем підтримки прийняття рішень у правовій сфері.

### Використана література

1. Т. Мото-Ока. ЭВМ пятого поколения. – М., Финансы и статистика, 1984.
2. А.Е. Якушин, Ф.И. Андон, В.А. Резниченко. Логические модели интеллектуальных систем. – К., 1999.
3. Д.А. Поспелов. Представление знаний в человеко-машинных и робототехнических системах. – М., 1984.
4. Е.В. Дмитриева, А.С. Клещеев. Формальное описание неагортмических языков программирования // Средства реализации систем искусственного интеллекта: Сб. науч. пр. – Владивосток, 1978.
5. Г.Р. Фирдман. Формализм представления глобального знания // Средства реализации систем искусственного интеллекта: Сб. науч. пр. – Владивосток, 1978.
6. А.И. Шевченко, В.А. Яценко. Может ли компьютер мыслить? // Искусственный интеллект. – 2005. – № 4.
7. Ю.С. Рамський, Логічні основи інформатики. – К., 2003.
8. М. Минский. Фреймы для представления знаний. Пер. с англ. – М., Энергия, 1979.

~~~~~ \* \* \* ~~~~~