УДК. 681.3

В.В. Литвин, Р.Б. Кравець
НУ "Львівська політехніка",
кафедра “Інформаційні системи та мережі”

ІНТЕЛЕКТУАЛЬНІ ІНФОРМАЦІЙНІ СИСТЕМИ З КІЛЬКОМА ІНТЕРПРЕТАТОРАМИ

© Литвин В.В., Кравець Р.Б., 2001.

This paper describes the concept of intelligent information systems with many . It considers main features of such systems. It describes an example of intelligent information system with many development in human resources planning.

У даній статті досліджується питання проектування інтелектуальних інформацій-них систем з кількома інтерпретаторами. Наводяться основні означення та задачі, які необхідно розв’язати при розробці таких систем.

Згідно з [1] основною характеристикою промислової програми є рівень її складності. Побудова моделей саме складних систем та моделей їх поведінки на основі змістовного опису об'єктів, цілей, обмежень, критеріїв, законів функціонування та інших елементів є однією з головних проблем сучасної теорії прийняття рішень [2].

Для вивчення складності систем проведено багато досліджень. Так, в роботі [3] сформульоване таке твердження.

Складні системи часто є ієрархічними і складаються із взаємозалежних підсистем, які у свою чергу також можуть бути розділені на підсистеми і т.д., аж до найнижчого рівня.

Ієрархія в системі передбачає певну "нерівноправність" – підпорядкування одних елементів системи іншим. Ієрархічні системи широко розповсюджені: наприклад, система зв'язку, система керування транспортом і багато інших завжди організовані за ієрархічним принципом, що дозволяє паралельно виконувати різні операції, працювати з окремими інформаційними масивами тощо. Необхідність ієрархічної організації в системах виникає тоді, коли централізоване опрацювання інформації або просто неможливе, або потребує таких затрат часу чи засобів, що не може реалізуватись через технічні умови [4].

Основна задача ієрархічної організації – розподіл функцій опрацювання інформації і прийняття рішення між окремими частинами системи (підсистемами).

У своїй роботі Саймон називає ієрархічні системи *розкладними*, якщо вони можуть бути розділені на окремі частини, і *майже розкладними*, якщо їх складові не є абсолютно незалежними. На основі цього можна сформулювати загальне твердження для всіх складних систем [5]:

Внутрішньокомпонентний зв'язок сильніший, ніж зв'язок між компонентами. Ця обставина дозволяє відокремлювати "високочастотні" взаємодії всередині підсистем від "низькочастотної" динаміки взаємодії між підсистемами.

Різниця між внутрішньокомпонентними і міжкомпонентними взаємодіями обумовлює розділення функцій між частинами системи і дає можливість відносно ізольовано вивчати кожну частину.

Розглянемо побудову складної інтелектуальної інформаційної системи (ІІС) $S_{ІІС}$=<R, B, I>, де R – база даних (БД); B – база знань (БЗ), яка містить множину модулів; I – інтерпретатор (машина логічного виведення, МЛВ), який в циклі виконує такі дії: визначає множину пар {модуль; набір поточних даних, на якому цей модуль задовольняється}, виконує відповідні дії, проводячи зміни в базі даних [6]. Нехай ІІС можна розбити на m підсистем S^i (i=1,2,…, m). Для кожної підсистеми S^i визначимо свою базу даних $R_i \subseteq R$, що містить інформацію про об'єкти цієї підсистеми.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Виходячи з означення інтелектуальної інформаційної системи, цілком природно для кожної підсистеми S^i (i=1,2,…, m) визначити свій власний інтерпретатор I_i. Але задачі, які опрацьовує підсистема, що знаходиться вище за ієрархією, залежать від результату прийнятого рішення у нижчих за ієрархією підсистемах (або навпаки), що веде до простоювання інтерпретаторів. Тобто m інтерпретаторів не потрібно. У той же час цікаво розглянути роботу системи, яка містить більше ніж один інтерпретатор.

Інтелектуальна інформаційна система з кількома інтерпретаторами - це інформаційна система, що містить два або більше інтерпретатори.

Тобто в основі такої системи лежить ідея використати для розв'язання однієї задачі кілька інтерпретаторів, що погоджено працюють. Розрахунок робиться на те, що якщо одній МЛВ для розв'язання задачі потрібно час t, то n МЛВ зможуть розв'язати цю задачу за час t/n. Однак це ідеальне прискорення вдається отримати лише в дуже спеціальних випадках. Тому метою є побудова алгоритмів функціонування ІІС з кількома інтерпретаторами, які можуть виходячи із кількості МЛВ, досягнути максимальної вигоди для конкретної задачі.

Розглянемо деяку ІІС з кількома інтерпретаторами I_1, I_2,…, I_n. Введемо основні означення, що характеризують такі системи.

Означення 1. Інтерпретатор I називається активним в момент часу t, якщо він працює у цей момент часу.

Множину активних інтерпретаторів у момент часу t позначатимемо $I_a(t)=\{I_{a_1(t)},...,I_{a_m(t)}\}$. Число $a_m(t)$ будемо називати степенем паралельності ІІС з кількома інтерпретаторами у момент часу t.

Означення 2. Потужністю ІІС з кількома інтерпретаторами P будемо називати максимальний степінь паралельності ІІС протягом всього часу розв'язування задачі:

$$P=\max_t a_m(t)\cdot$$

Означення 3. Інтерпретатор I називається пасивним в момент часу t, якщо він не працює у цей момент часу.

Число пасивних інтерпретаторів у момент часу t є $b_k(t)=n-a_n(t)$.

Означення 4. Інтерпретатор I називається бездіяльним, якщо він є пасивним протягом всього часу розв'язування задачі.

Потужність служить основною характеристикою ІІС з кількома інтерпретаторами. Вона визначає максимальну кількість інтерпретаторів, серед яких немає бездіяльних.

Означення 5. Прискоренням ІІС з кількома інтерпретаторами називається відношення:

$$S_n=\frac{\text{час виконання алгоритму з 1 МЛВ}}{\text{час виконання алгоритму з n МЛВ}}$$

Означення 6. Ефективністю ІІС з кількома інтерпретаторами називається величина

$$E_n=\frac{S_n}{n}\cdot$$

Розглянемо формальну модель прискорення, в якій

$$S_n=\frac{T_1}{(a_1+a_2/k+a_3/n)T_1+t_d},$$

де T_1 - час, який затрачається однією МЛВ для розв'язування задачі;

a_1 - частина операцій, що виконується тільки однією МЛВ;

a_2 - частина операцій, що виконуються із середнім степенем паралелізму k<n;

a_3 - частина операцій, що виконується із степенем паралелізму n;

t_d - загальний час, який потрібний на підготовку даних.

Випадок 1. $a_1=a_2=0$; $a_3=1$; $t_d=0$. Тоді $S_n=n$ і прискорення є максимальне.

Випадок 2. $a_1=a_3=0$; $a_2=1$; $t_d=0$. Тоді $S_n=k<n$ і прискорення дорівнює середньому степеневі паралелізму.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

Випадок 3. $a_2=0$; $t_d=0$; $a_1=\alpha$; $a_3=1-\alpha$. Тоді

$$S_n = \frac{1}{\alpha+(1-\alpha)/n}.$$

Остання формула виражає закон Уера-Амдаля [7]. Це означає, що кожна операція відбувається або з максимальним, або з мінімальним паралелізмом, і затримки відсутні.

Нехай в деякій задачі половина операцій може виконуватися паралельно, а половина ні. Тоді $\alpha=1/2$ і $S_n = \frac{2}{1+n^{-1}} < 2$. Тобто прискорення менше за два.

Випадок 4. t_d велике. Тоді можна отримати $S_n<1$. Цей варіант відображає випадок, коли використання кількох МЛВ є менш вигідним, ніж однієї МЛВ.

Однією з головних цілей при конструюванні паралельного алгоритму є досягнення якомога більшого прискорення і в ідеальному випадку $S_n=n$. Однак цього не завжди можна досягнути. Максимальне прискорення можна отримати тільки для тривіальних задач. Головні фактори, що обумовлюють відхилення від максимального прискорення, такі:

- відсутність максимальної паралельності в алгоритмі і/або незбалансованість навантаження МЛВ;
- обмін даних, конфлікти і час синхронізації.

Синхронізація необхідна в тих випадках, коли деякі відрізки обчислень повинні бути закінчені скоріше, ніж зможе продовжитись процес в цілому.

Більшість алгоритмів являють собою суміш фрагментів з трьома різними степенями паралельності, які ми будемо називати максимальним, частковим і мінімальним паралелізмом. Останній термін відноситься до тих фрагментів алгоритму, де можна використати лише одну МЛВ. Але якщо навіть алгоритм сам по собі володіє максимальним паралелізмом, ситуація для цієї паралельної системи може ускладнюватися проблемою балансування навантаження. Під балансуванням навантаження розуміємо такий розподіл завдань між МЛВ системи, які дозволяють задіяти кожну МЛВ корисною роботою якомога більшу частину часу.

Балансування навантаження може здійснюватись як статично, так і динамічно. При статичному балансуванні завдання розподіляються між МЛВ до початку розв'язування задачі. При динамічному балансуванні завдання розподіляються між МЛВ під час алгоритмічного процесу. Корисним поняттям, яке пов'язане із динамічним балансуванням навантаження, є поняття банку завдань. МЛВ отримують з банку чергові завдання, які готові до виконання.

Досягнення мети функціонування системи будемо називати глобальним розв'язком задачі. На основі цієї мети сформуємо простір цілей, інакше кажучи множину підзадач, з яких складається основна задача системи.

На наш погляд, найкращим є варіант, коли одна МЛВ розв'язує одну окрему підзадачу, тому що на передачу даних (множина попередніх станів E, множина зазначених пар Z) між МЛВ затрачається певний час. А коли починається розв'язування нової підзадачі, попередні дані не потрібні. З іншого боку при такому розподілі система набуває певної гнучкості, тому що керуючу функцію u інтерпретатора можна визначити трьома різними способами, залежно від числа змінних [6]. Це дає змогу для розв'язування задачі використовувати в системі різні рівні інтелектуальності, тим самим керувати процесом пошуку глобального розв'язку, синхронізацією та балансуванням навантаження.

Розглянемо приклад побудови ІІС з двома інтерпретаторами. За проблемну галузь виберемо діяльність служби працевлаштування. Ми вважаємо, що сучасна служба працевлаштування повинна мати у своїй основі автоматизовану інформаційну систему, яка містить інформацію про резерви та потреби у кваліфікованих кадрах та надає послуги з працевлаштування для максимальної реалізації професійного рівня, вимог до оплати праці осіб, що шукають роботу.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

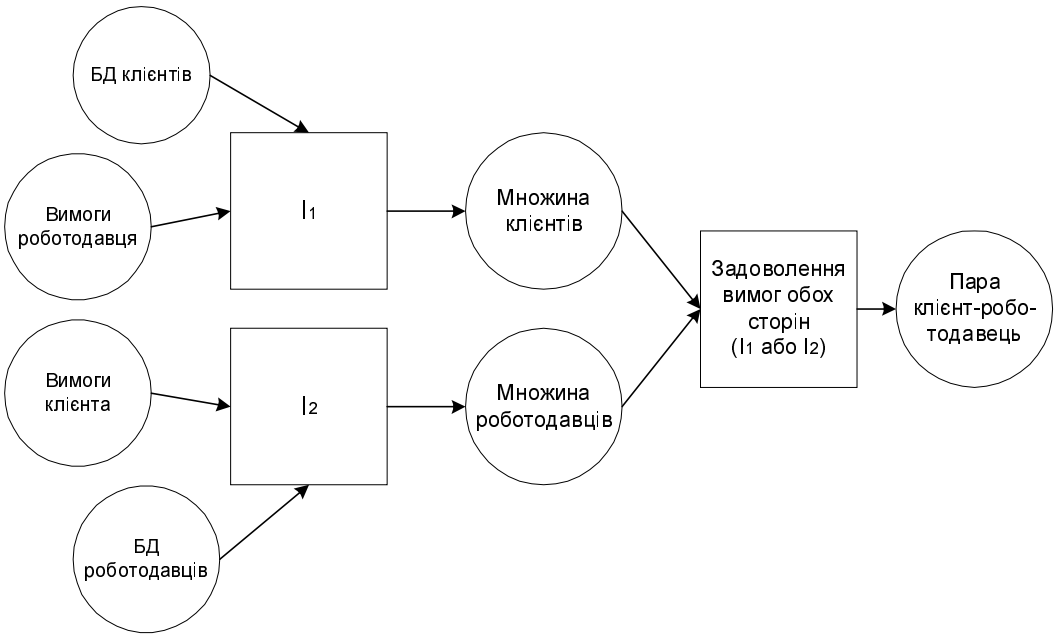

Рис.1. *Функціонування ІІС служби працевлаштування*

Службу працевлаштування можна розглядати як комерційного посередника, який шукає партнерів у здійсненні акту купівлі-продажу продукту, що ним є інформація про робочу силу та робочі місця. Тут покупцем виступає роботодавець, що пропонує роботу, а продавцем — клієнт, що пропонує свою робочу силу, або вони міняються ролями. Служба працевлаштування прагне задовольнити вимоги обох сторін та зацікавлена у найповнішому врахуванні побажань своїх клієнтів. Отже, у формуванні відносин отримання роботи діють три учасники, кожний з яких має власні інтереси: роботодавці *R*, претенденти на отримання роботи (клієнти) *K* та служба працевлаштування *S* [8].

Глобальну мету функціонування служби працевлаштування можна розбити на дві окремі підмети (підзадачі):

- задоволення вимог роботодавця;
- задоволення вимог клієнта,
- які можна розв'язувати паралельно. Тому визначимо для системи два інтерпретатори I_1, I_2, які б паралельно розв'язували підзадачі 1 та 2. Після їх розв'язання один із інтерпретаторів можна використати для розв'язання основної задачі – задоволення вимог обох сторін.

Процес функціонування системи зображено на рисунку.

На початку роботи ІІС перший інтепретатор I_1 використовує інформацію про вимоги роботодавців та про знання клієнтів. Він проводить відбір тих клієнтів, що задовольняють вимоги роботодавця. Другий I_2 використовує інформацію про вимоги клієнтів та дані про роботу. Його завдання полягає у відборі тих роботодавців, що задовольняють вимоги клієнтів. Для розв'язання останньої задачі використовуються розв'язки двох попередніх підзадач. Процес знаходження кінцевого розв'язку може здійснити будь-який інтерпретатор. З точки зору мінімізації передачі даних краще використати той, що сформував більшу множину даних.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

На завершення сформулюємо основні переваги ІІС з кількома інтерпретаторами порівняно з інтелектуальними системами, що містять одну МЛВ:

- прискорення часу розв'язання задачі;
- можливість використання різних функцій керування залежно від підзадачі;
- використання в одній системі кількох методів представлення знань;
- перевірка роботи системи заміною одних інтерпретаторів на інші;
- збільшення надійності отриманого розв'язку;
- координація роботи системи загалом.

1. Буч Г. Объектно-ориентированный анализ и проектирование. - М.: Издательство Бином, СПб: Невский диалект, 1998. 2. Клыков Ю.И., Горьков Л.Н. Банки данных для принятия решений. - М.: Советское радио, 1980. 3. Courtois, P. On time and Space Decomposition of Complex Structures. Communications of the ACM vol. 28(6), - 1985. - p. 596. 4. Моисеев Н.Н. Математические задачи системного анализа. - М.: Наука, 1981. 5. Simon, H. The Sciences of the Artificial. Cambridge, MA: The MIT Press, 1982. - p. 218 6. Литвин В.В. Формальна модель інтелектуальної інформаційної системи // Вісн. ДУ "Львівська політехніка". - 1999. - № 386. - С. 104-108. 7. Дж. Ортега "Введение в паралельные и векторные методы решения линейных систем". - М.: Мир, 1991. 8. Литвин В.В., Нікольський Ю.В. Застосування методів логічного програмування для автоматизації процедур прийняття рішень у діяльності служби працевлаштування // Вісн. ДУ "Львівська політехніка". - 1998. - № 330 - С. 153-163.

УДК 683.1

Е. Марецька*, Р.Б. Кравець
* Інститут інформатики та управління,
НУ "Львівська політехніка", кафедра "Інформаційні системи та мережі"

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ У СФЕРІ СОЦІАЛЬНОГО СТРАХУВАННЯ

© Марецька Е., Кравець Р., 2002

This paper considers classification method for knowledge discovery in databases. The induction of decision tree algorithms for supervised classification is described. Authors propose to use this method to predict the result of sport match.

Стаття містить класифікаційні методи видобування знань у базах даних. Описаний метод індукції у деревах рішень. Автори пропонують використовувати ці методи для передбачення результату спортивних матчів

1. Вступ

Інформаційні технології все глибше проникають у різні сфери економічної діяльності. Однією з таких галузей є соціальне страхування. Для ефективного ведення страхового бізнесу доводиться фіксувати і аналізувати велику кількість параметрів, які впливають на виникнення страхових випадків. При цьому страховим компаніям доводиться враховувати усі ці фактори і приймати рішення, спрямовані на покращання якості виконання обов'язків перед клієнтами і водночас на отримання максимального прибутку.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua