

УДК 519.68

¹Ю.М. Тесля, ²Т.В. Савельєва, ¹О.М. Курілко, ¹Ю.О. Остапчук, ¹П.В. Каюк¹Київський національний університет будівництва і архітектури, Київ²Черкаський державний технологічний університет, Черкаси

ЗАСТОСУВАННЯ РЕФЛЕКТОРНИХ СИСТЕМ ОБРОБКИ ПРИРОДНОМОВНОЇ ІНФОРМАЦІЇ В УПРАВЛІННІ ПРОЕКТАМИ

Розглянуто проблему створення на базі математичного апарату теорії несилової взаємодії моделей функціонування систем обробки природномовної інформації. Класифіковано задачі, які зможуть вирішувати такі системи в управлінні проектами.

Ключові слова: природномовна система, управління проектами, теорія несилової взаємодії

Постановка проблеми

Без застосування сучасних засобів обробки природномовної інформації неможливо розраховувати на зручність та ефективність застосування різноманітних інформаційних систем і технологій управління проектами. Зрозуміло, що в такій складній сфері як управління проектами, де діяльність менеджерів і спеціалістів пов'язана з обробкою значної кількості даних, неможливо обійтись без застосування засобів обробки природномовної інформації. Але сучасний український ринок інформаційних технологій (ІТ) не пропонує компаніям спеціалізовані інтелектуальні системи обробки природномовної інформації в управлінні проектами.

Тому виникла необхідність створення проблемно - орієнтованих інтелектуальних систем, зданих з природномовного представлення різноманітних категорій проектів «відібрати» необхідну для менеджерів і спеціалістів інформацію та застосовувати її для вирішення задач планування, управління вартістю, управління ризиками та ін.

Аналіз останніх досліджень

Є чимало досліджень з питань побудови систем обробки природної мови [1-6]. Сучасні системи обробки природномовної інформації, що виникли на перетині комп'ютерної лінгвістики та інформаційно-пошукових систем [1], являють собою системи, що слабо пристосовані до вирішення задач формалізації управлінської інформації.

Сучасний український ринок інформаційних технологій (ІТ) не пропонує компаніям інструментальні інтелектуальні системи, які можна було б з найменшими витратами застосовувати для обробки природномовних текстів в управлінні проектами. Відомі вітчизняні та зарубіжні системи

обробки природномовної інформації є експериментальними або вузькоспеціалізованими і не задовольняють вимоги систем автоматизованої переробки інформації в управлінні проектами. Крім цього, витрати на створення таких систем є значними, що не дозволяє їх широко використовувати на більшості проектно-орієнтованих підприємств [1-5].

Але тим не менш, без таких засобів значно знижується продуктивність автоматизованих систем при наповненні інформаційної бази управління проектами, підвищуються витрати на виконання рутинних дій, таких як формальне представлення і введення інформації, пошук в інформаційній базі та ін. [7].

Формулювання цілей статті

Використання систем обробки природномовної мови в управлінні проектами пов'язано з низкою специфічних особливостей цієї галузі діяльності, значною тривалістю та складністю процесу створення продукту проекту, значними змінами в документації по ходу реалізації проектів, неповнотою документації, низькою надійністю джерел інформації, різноманіттям вихідної документації, що формується за запитами різних користувачів, важкою формалізацією більшості вихідних документів [7]. Відповідно до цього, вимоги та обмеження до систем обробки природномовних текстів визначаються мобільністю, надійністю, високою продуктивністю, простотою зміни інформаційної бази, децентралізацією процесу обробки даних.

Метою роботи є розробка з використанням математичного апарату теорії несилової взаємодії моделей, методу та засобів формалізації природномовної інформації, на базі яких можна

було б створювати економічно вигідні, ефективні та прості системи управління проектами.

Основний матеріал дослідження

Ефективність побудови засобів автоматизованого управління проектами багато в чому залежить від мовних засобів описання об'єктів та процесів управління [7]. Тому, адаптивність та мобільність мов спілкування з «комп'ютером», непроцедурний характер опису об'єктів та процесів управління, використання «контекстів» та «замовчуванням» – неодмінна умова раціональної конструкції мовних засобів. Ці вимоги мають суперечливий характер, обумовлений орієнтацією на непрофесійних користувачів, з одного боку, і обмеженими можливостями (як теоретичними, так і технічними) з іншого. Вирішення даного протиріччя пов'язано:

а) з поділом вхідних інформаційних потоків на ті, що формалізуються, (і, відповідно, вимагають для своєї обробки традиційних діалогових програмних засобів) і ті, що важко формалізуються, (вимагають значних витрат на формалізацію);

б) з використанням результатів отриманих в теорії несилової взаємодії для побудови засобів обробки природно-мовної інформації [6].

Формалізуємо задачу застосування систем обробки природномовної інформації для прийняття рішень в проектах. Ця задача формулюється так: необхідно виконати такі дії (чи прийняти ті чи інші рішення) в проекті, які відповідають семантико-прагматичному наповненню природномовного тексту.

Для вирішення цієї задачі, і, виходячи з інформаційної надлишковості природної мови, розкладемо вхідний текст на кластери, серед яких зазвичай будуть знаходитися ті, що і визначають необхідну за даних умов реакцію. Оскільки знати заздалегідь, які це кластери неможливо, сформуємо ряд комбінаторних регулярностей вхідного тексту і в процесі навчання відберемо ті, які найбільш впливають на рішення. Це визначається, виходячи з такої моделі:

1. Множина елементів природної мови, що впливають на рішення, розкладається на кластери, які містять параметри:

$$X = \{x_i\}, i = \overline{1, q},$$

де X – вхідний текст; x_i – кластер, що впливає на рішення; q – кількість кластерів.

2. За частотними характеристиками можна визначити приблизні значення імовірності кожного з рішень (кожної з дій) за умови, що зустрічається той чи інший кластер:

$\forall R_j \in R, x_i \in X : p(R_j / x_i) \approx n(R_j / x_i)$, де $n(R_j / x_i)$ – частота дії (рішення) R_j за умови, що кластер x_i був наявний у вхідному тексті; $p(R_j / x_i)$ – імовірність дії (рішення) R_j за умови, що кластер x_i був наявний.

3. Для кожної дії (рішення) R_j за безумовної імовірності $p(R_j)$ і окремим умовним ймовірностям $p(R_j / x_1), \dots, p(R_j / x_i), \dots, p(R_j / x_q)$ необхідно оцінити спільну умовну імовірність цього рішення (дії) $p(R_j / x_1, \dots, x_i, \dots, x_q) = p(R_j / X)$.

Оскільки в даній інтерпретації, процес прийняття рішення можна представити як процес визначення найбільш значних інформаційних впливів вхідного тексту на множину дій (рішень), то для вирішення цієї задачі використаємо підхід, що базується на математичному апараті теорії несилової взаємодії. Мірою оцінки ситуації, відносно рішення в теорії несилової взаємодії, є визначеність і інформованість. Зв'язок між визначеністю та імовірністю задається формулою:

$$d = \begin{cases} 0,5 \cdot \sqrt{\frac{p}{1-p} + \frac{1-p}{p}} - 2, \text{ при } p \geq 0,5 \\ -0,5 \cdot \sqrt{\frac{p}{1-p} + \frac{1-p}{p}} - 2, \text{ при } p < 0,5 \end{cases}$$

де p – імовірність рішення; d – визначеність цього рішення.

Зв'язок між інформованістю та імовірністю задається формулою:

$$i = \frac{0,5}{\sqrt{p \cdot (1-p)}}.$$

Міра несилової дії $b_j \in B, j = \overline{1, n}$ повинна відбивати різницю у визначеності і інформованості "до" і "після" змін в предметній області. І саме зміна визначеності і інформованості є причиною зміни імовірності

$$d_0 \rightarrow d_j \wedge i_0 \rightarrow i_j \Rightarrow p_0 \rightarrow p_j,$$

де p_0 – імовірність рішення (дії) R_0 ; d_0 – визначеність рішення (дії) R_0 ; i_0 – інформованість відносно рішення (дії) R_0 ; d_j – визначеність рішення (дії) R_0 , яка сформована кластером вхідного тексту $b_j \in B$; i_j – інформованість, яка відповідає визначеності d_j ; p_j –

імовірність рішення (дії) R_0 задана всіма кластерами $b_j \in B$.

При цьому, з [6]:

$$d_0 \pm \frac{1}{2} \sqrt{\frac{p_0}{1-p_0} + \frac{1-p_0}{p_0} - 2}; i_0 = \frac{0,5}{\sqrt{p_0 \cdot (1-p_0)}}; \quad (1)$$

$$d_j \pm \frac{1}{2} \sqrt{\frac{p_j}{1-p_j} + \frac{1-p_j}{p_j} - 2}; i_j = \frac{0,5}{\sqrt{p_j \cdot (1-p_j)}}. \quad (2)$$

Необхідно знайти перехід від різниці у визначеності рішення (дії) R_0 у випадку, коли в тексті присутній кластер $b_j \in B$ до визначеності прийняття рішення R_0 що задається всіма кластерами вхідного тексту [6]

$$\forall d_j, i_j : d_0 \rightarrow d_\Sigma \wedge i_0 \rightarrow i_\Sigma \Rightarrow p_\Sigma = 0,5 + \frac{d_\Sigma}{2_\Sigma},$$

де d_Σ – визначеність рішення R_0 , яка сформована всіма кластерами $b_j \in B$; i_Σ інформованість, що відповідає визначеності d_Σ ; p_Σ – імовірність рішення R_0 при впливі всіх кластерів, що входять у вхідний текст B .

Для вирішення цієї задачі застосуємо математичний апарат теорії несилової взаємодії. В концепції побудови рефлекторних інтелектуальних систем, яка заснована на математичному апараті теорії несилової взаємодії, реалізуються принципи універсального моделювання закономірностей у заданій предметній області, що статистично проявляються. В основі функціонування рефлекторних інтелектуальних систем лежить метод вироблення адекватної всім зовнішнім впливам реакції інтелектуальної системи. Цей метод забезпечує вибір найбільш імовірної реакції інтелектуальної системи на множину вхідних впливів, при відомих ймовірностях вибору реакції на кожний вхідний вплив, а також на деякі комбінації вхідних впливів.

Рефлекторний метод прямо розраховує адекватну вхідним впливам реакцію інтелектуальної системи (рис.1). Адекватність реакції підтверджена в експериментальних дослідженнях [6]. Рефлекторні інтелектуальні системи вирішують різноманітні задачі в областях інтелектуальної діяльності людини [8-9].

Метод розрахунку адекватної сенсу вхідного тексту реакції в управленні проектами має етапи:

1. Розрахунок визначеності реакцій в проекті відносно впливів кластерів тексту. З (1)-(2):

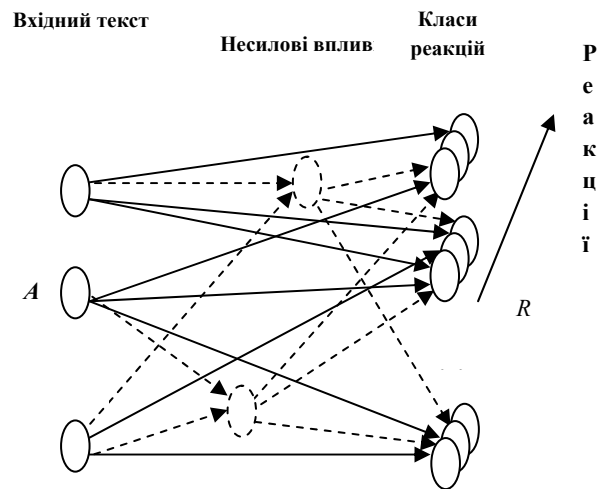
$$d(R_i) = \pm 0,5 \sqrt{\frac{p(R_i)}{1-p(R_i)} + \frac{1-p(R_i)}{p(R_i)} - 2};$$

$$i(R_i) = \sqrt{d^2(R_i) + 1};$$

$$d(R_i / A_j) = \pm 0,5 \sqrt{\frac{p(R_i / A_j)}{1-p(R_i / A_j)} + \frac{1-p(R_i / A_j)}{p(R_i / A_j)} - 2};$$

$$i(R_i / A_j) = \sqrt{d^2(R_i / A_j) + 1},$$

де $p(R_i)$ – безумовна імовірність $p(R_i)$ вибору рішення (дії) R_i ; $d(R_i)$ – визначеність в проекті відносно рішення (дії) R_i ; $i(R_i)$ – інформованість в проекті відносно рішення (дії) R_i ; $p(R_i/A_j)$ – умовна імовірність вибору рішення (дії) R_i за наявності у вхідному тексті кластера A_j ; $d(R_i/A_j)$ – визначеність в проекті відносно рішення (дії) R_i , за наявності у вхідному тексті кластера A_j ; $i(R_i/A_j)$ – інформованість в проекті відносно рішення (дії) R_i , при наявності у вхідному тексті кластера A_j .



Примітка: пунктиром відмічені віртуальні об'єкти, які відображують спільність дії (синергетичний ефект)

Рис.1. Схема розрахунку величини впливу в рефлекторних інтелектуальних системах

2. Розрахунок додаткової імовірності, яку привносять кластери вхідного тексту в інтелектуальну систему:

$$\Delta d(R_i / A_j) = d(R_i / A_j) \cdot i(R_i) - d(R_i) \cdot i(R_i / A_j),$$

де $\Delta d(R_i/A_j)$ – зміна визначеності рішення (дії) R_i при наявності у вхідному тексті кластера A_j .

3. Розрахунок сумарного впливу на менеджерів, що приймають рішення чи спеціалістів, які виконують якісь дії:

$$\Delta d_\Sigma(R_i) = \sum_j \Delta d(R_i / A_j),$$

де $\Delta d_\Sigma(R_i)$ – зміна визначеності рішення (дії) R_i .

Из (5.10):

$$\Delta i_\Sigma(R_i) = \sqrt{\Delta d_\Sigma^2(R_i) + 1},$$

де $\Delta i_\Sigma(R_i)$ – зміна інформованості в проекті, відносно рішення (дії) R_i .

4. Розрахунок нової визначеності рішень (дій) в проекті:

$$\overline{d}(R_i) = \Delta d_{\Sigma}(R_i) \cdot i(R_i) + d(R_i) \cdot \Delta i_{\Sigma}(R_i),$$

де $\overline{d}(R_i)$ – нова визначеність ставлення менеджменту проекту до рішення (дії) R_i .

5. Вибір максимально визначеного рішення (дії) R_k :

$$R_k : R_k \in R \wedge \overline{d}(R_k) = \max_i (\overline{d}(R_i)).$$

Під час створення рефлексорних систем обробки природномовної інформації з використанням наведеного методу, необхідно вирішувати дві головні задачі: виділити у вхідному тексті та формалізувати кластери, що впливають найбільше, а також реакції менеджменту проекту з накопиченням статистичної інформації про зв'язки між ними; адаптувати рефлексорний алгоритм до особливостей природної мови.

На основі запропонованої моделі розпочата розробка рефлексорної системи обробки природномовної інформації – системи дискретної обробки семантики текстів в управлінні проектами (ДОСТУП).

Основна ідея, що лежить в основі цієї системи, полягає в ідентифікації алгоритму та основних параметрів запиту по кластерам. природномовних текстів, що надходять. В основі системи ДОСТУП лежить математична модель несилової взаємодії в об'єктах та процесах, що формують природномовні тексти. Система ДОСТУП буде забезпечувати формування результуючого, формального представлення семантичної складової вхідного тексту за її представлення засобами природної мови.

Наповнення системи ДОСТУП буде формуватися сукупністю довільних кластерів вхідних текстів, опису реакцій та статистичних зв'язків між кластерами і реакціями. Вибір найбільш інформативних кластерів вхідних текстів буде виконуватися в процесі навчання системи ДОСТУП.

Під час обробки вхідної інформації в системі ДОСТУП засоби природномовного спілкування будуть використовуватися для:

- ідентифікації алгоритму вирішення задачі, про яку йде мова в природномовному тексті;
- визначення об'єкта та суб'єктів управління.
- встановлення часового інтервалу;
- формального представлення параметрів задачі, що розв'язуються;
- представлення в інформаційній базі компонентів управління проектами.

За допомогою системи ДОСТУП будуть вирішуватися задачі:

- Формалізації природномовного тексту чи його частини;

- Відбору необхідної інформації з бази даних інформаційної системи управління проектами та її представлення в зручному для користувача вигляді;

- Внесення змін в інформаційне представлення компонентів проектів (в базі даних). Зокрема параметрів робіт, ризиків, вартості, виконавців, відповідальних і т.п.;

Система буде реалізовувати три форми взаємодії з користувачем це обробки:

- тексту, який представлено в інформаційній базі чи отримано з зовнішніх джерел (наприклад, з мережі Internet);

- звертання користувача, який набирається на клавіатурі.

- звертання користувача на усній мові.

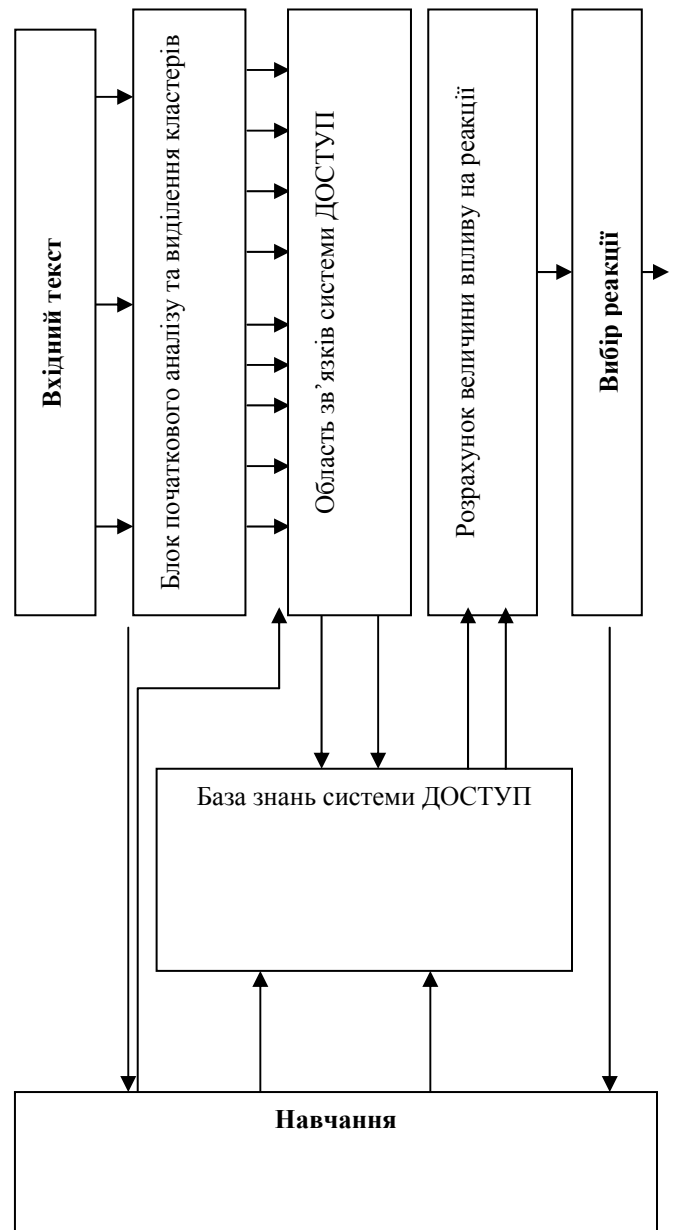


Рис.2. Схема інформаційних потоків в системі ДОСТУП

В основі системи ДОСТУП знаходиться рефлекторна модель поведінки в середовищі функціонування, що неоднозначно трактується та містить комплекс програм, які забезпечують інтерфейс між користувачем, інформаційною базою, засобами наповнення інформаційної бази, інформаційно-пошуковою системою і системою управління проектами.

Схема інформаційних потоків в системі ДОСТУП показана на рис.2.

Зазначимо, що в системі ДОСТУП йдеться не про «розуміння» обмеженої природної мови, а про систему, що сприймає інформаційні представлення на природній мові без обмежень.

Висновки та перспективи подальших досліджень

В процесі досліджень було показано, що математичний апарат теорії несилової взаємодії досить простий і зручний для побудови систем обробки природномовних текстів. Розроблено математичну модель, яка буде знаходитися в основі системи дискретної обробки семантики текстів в управлінні проектами (система ДОСТУП). В подальшій роботі планується розробити структуру та алгоритми системи ДОСТУП та провести її дослідну перевірку в виробничих умовах для управління проектами.

Список літератури

1. Попов Э.В. Общение с ЭВМ на естественном языке / Э.В. Попов. - М.: Наука, 1982. - 360 с.
2. Бондаренко М.Ф. Основы теории синтеза удивительных структур мовних систем штучного інтелекту / М.Ф. Бондаренко ХДТУРЕ. - К.: 1997. - 264 с.
3. Тесля Ю.М. Застосування теорії інформаційної взаємодії до побудови систем класифікації образів // Праці сьомої міжнародної конференції "Укробраз 98", Київ, 26-30 жовтня 1998р. с.122-123.
4. Тесля Ю.М., Застосування теорії інформаційної взаємодії до побудови систем смислового пошуку інформації / Ю.М. Тесля, О.М. Косенкова // Радиоэлектроника и информатика, 2000. - №1. - С.103-105.
5. Файн В.С. Распознавание образов и машинное понимание естественного языка / В.С. Файн. - М.: Наука, 1987. - 173 с.
6. Тесля Ю.Н. Введение в информатику природы / Ю.М. Тесля // К.: Маклаут, 2010. - 255 с.
7. Тесля Ю.Н., Тимченко А.А. Опыт разработки и применения в строительстве инструментальных программных средств естественно-языкового общения / Ю.Н. Тесля, А.А. Тимченко // Тезисы докладов на II международном научно-техническом семинаре

"Теоретические и прикладные проблемы моделирования предметных областей системах баз данных и знаний", Туансе, 20-26 сентября 1993г. Киев: МГП "Тираж". С.226-228.

8. Олексієнко М.М. Метод прогнозування чисельності захворювань від впливу шкідливих речовин, який базується на моделі несилової взаємодії / М.М. Олексієнко // «Східно-європейський журнал передових технологій». - Харків. - 2009. - №1. - С.34-38.

9. Тесля Ю.М. Застосування рефлекторного підходу до побудови інтелектуальних систем оцінювання інвестиційних пропозицій / Ю.М. Тесля, П.В. Каюк, М.Л. Чернова // - Всеук-раїнський збірник наукових праць: Гірничі, будівельні, дорожні та меліоративні машини. - Видавництво КНУБА. №73-2009. - С.82-87.

Стаття надійшла до редколегії 2.02.2011

Рецензент: д-р техн. наук, доц. Н.С. Бушуєва Київський національний університет будівництва і архітектури, Київ