

Ромащенко Р.А., викладач  
Академії муніципального  
управління, м. Київ

## АНАЛІЗ ІСНУЮЧИХ МОДЕЛЕЙ НАДАННЯ ЗНАНЬ ДЛЯ ВИКОРИСТАННЯ В СИСТЕМАХ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

---

*Ціллю статті є дослідження та аналіз існуючих методів надання знань в системах підтримки прийняття рішень. Були визначені функції підсистеми надання знань в СППР. Для належного виконання наведених функцій моделі надання знань повинні задовольняти певним вимогам, що були визначені в статті.*

*Целью статьи является исследование и анализ существующих методов предоставления знаний в системах поддержки принятия решений. Были определены функции подсистемы предоставления знаний в СППР. Для надлежащего выполнения приведенных функций модели предоставления знаний должны удовлетворять определенным требованиям, что были определены в статье.*

*Research and analysis of existent methods of grant of knowledge in the systems of support of acceptance of decisions is the purpose of the article. The functions of subsystem of grant of knowledge in SPPR were defined. For the proper implementation of the resulted functions of grant model knowledge must satisfy to the definite requirements, that was defined in the article.*

Стаття присвячена проблемам вибору моделі надання знань для використання в системах підтримки прийняття рішень та являє собою стислий огляд існуючих моделей з аналізом їхніх переваг та недоліків.

**Постановка завдання.** Автоматизація процесів прийняття рішень у складних організаційно-технічних системах шляхом створення автоматизованих систем підтримки прийняття рішень (СППР) є одним із актуальних напрямів науково-практичних досліджень. СППР являють собою людино-машинні об'єкти, що дозволяють особам, які приймають

рішення, використовувати дані, знання, об'єктивні й суб'єктивні моделі для аналізу та вирішення слабоструктурованих і неструктурованих проблем. Вирішення задач генерації та оцінки рішень у СППР базується на таких методах: дослідження операцій; моделювання, орієнтованих на знання. Тому постає проблема вдалого вибору моделі надання знань

**Аналіз останніх досліджень і публікацій, в яких започатковано вирішення даної проблеми.** Моделі надання знань вивчалися дослідниками Е.А.Трахтенгерцом[1], Т.А. Гавриловою[3], А.В. Назаровим [4], І.Б.Сероджою [5] та ін. Системи підтримки прийняття рішень глибоко вивчали Б.М, Герасимов, А.В. Тарасов, І.В.Токарев [7].

**Мета статті.** Проаналізувати існуючі моделі надання знань в системах підтримки прийняття рішень.

### **Викладення основного матеріалу.**

Застосування технологій обробки знань у СППР, за деякими прогнозами, являє собою найбільш перспективний напрямок їхнього розвитку. Хоча прогнози кінця минулого століття щодо швидкого створення повноцінних штучних інтелектуальних систем не виправдались, проте слід відзначити значні досягнення в цій сфері, зокрема: досліджено та формалізовано класичні моделі надання знань, розвиваються нові моделі, формалізовано значну кількість методів і методик добування знань, отримала значний розвиток теорія нечітких множин, штучних нейронних мереж, генетичних алгоритмів тощо.

У процесі проектування СППР на основі технологій обробки знань необхідно найбільш повно враховувати характерні особливості предметної області та забезпечувати адекватне надання знань щодо об'єктів прийняття рішень у базі знань СППР. Адекватність надання знань задачам, що вирішуються СППР, забезпечується правильним вибором моделі надання знань і відповідних методів логічного виводу. У даній статті проведено порівняльний аналіз властивостей відомих моделей надання знань з точки зору вимог, що до них висуваються, з метою здійснення обґрунтованого вибору моделі під час проектування інтелектуальних СППР.

Класичними моделями, що використовуються для надання знань у СППР, є семантичні мережі, фрейми, логічні та продукційні моделі [1]. Значного поширення набуло використання нейронних мереж, квантових і гібридних моделей надання знань.

Семантичні мережі є однією з найбільш поширених форм надання знань. Вони були запропоновані в 60-і роки минулого сторіччя як відображення уявлень про понятійний апарат людини. Значну роль у їхньому поширенні відіграли великі можливості семантичних мереж щодо

відображення довільної структури понять предметної області та наявність розроблених математичних методів вирішення задач на графах.

У загальному випадку семантичну мережу може бути зображено у вигляді графу  $G=(X, R)$ , де  $X=\{x_i\}$ ,  $i=\overline{1, n}$  – множина вершин графу,  $R=\{<x_i, x_j>\}$ ,  $i, j=\overline{1, n}$  – множина дуг [7]. Вершини графу відповідають поняттям предметної області, а дуги – відношенням між ними (рис. 1).

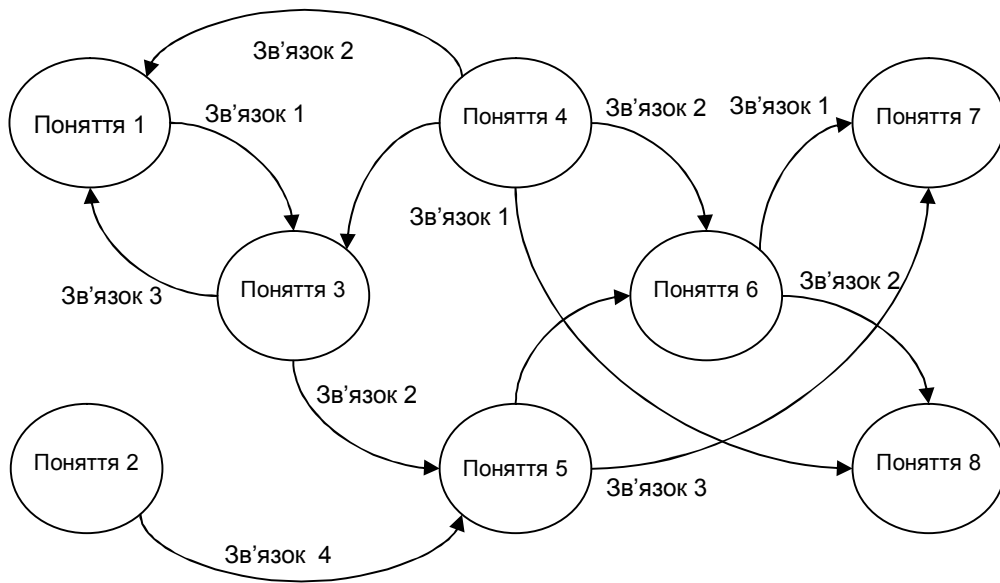


Рис. 1. Структура семантичної мережі

Як додаткові характеристики графу, залежно від реалізації, можуть бути введені типи вершин, дуг, імена і вагові коефіцієнти відношень тощо. Моделі на основі семантичних мереж застосовуються для вирішення таких актуальних задач, як побудова онтологій предметної області, побудова семантичного образу тексту та ін.

Основними перевагами семантичних мереж є універсальність і можливість відображення теоретично необмеженої кількості понять і відношень предметної області. Семантичні мережі придатні для масштабування, в них ефективно реалізується пояснення рішень, а експорт/імпорт знань є можливим практично в будь-які інші типи моделей надання знань.

Недоліки семантичних мереж – висока обчислювальна складність реалізації процедур пошуку і логічного виводу на насичених семантичних мережах.

Фреймові моделі знань були запропоновані у 70-і роки минулого сторіччя як окремий випадок семантичних мереж для відображення стереотипних ситуацій предметної області [4]. Фреймові моделі

застосовувались з метою спрощення процедур логічного виводу за рахунок типізації вершин і зв'язків у мережі. Один із головних принципів, закладених в основу фреймових моделей, – ієрархічний характер з ослабленими відношеннями наслідування. Типізація фреймів здійснюється шляхом створення фреймів-прототипів, властивості яких наслідуються фреймами-екземплярами. Фреймова модель дає можливість поєднати в одній структурі фрейму декларативні та процедурні знання за рахунок суміщення фактографічних знань і відповідних процедур їхньої обробки. Структуру фрейму-прототипу  $T_i^F$  можна зобразити у вигляді:

$$T_i^F = \{ \langle \text{ім'я фрейму} \rangle, \\ \{ \langle \text{ім'я слоту } 1 \rangle; \langle a_1, S_1, T_1 \rangle \dots \langle a_h, S_h, T_h \rangle; \langle H_1 \dots H_j \rangle; \langle q_1 \dots q_k \rangle; \langle D \rangle \}; \\ \{ \langle \text{ім'я слоту } n \rangle; \langle a_1, S_1, T_1 \rangle \dots \langle a_l, S_l, T_l \rangle; \langle H_1 \dots H_m \rangle; \langle q_1 \dots q_t \rangle; \langle D \rangle \} \},$$

де  $a$  – імена атрибутів, характерних для слоту,  $S$  – значення атрибутів,  $T$  – типи даних атрибутів;  $H$  – покажчики наслідування,  $q$  – зв'язок з іншими слотами або фреймами,  $D$  – покажчик процедури обробки.

Мережу фреймів можна зобразити у вигляді графу  $G=(F, R)$ , де  $F=\{f_i, f_i \in T^F\}$ ,  $i = \overline{1, n}$  – множина вершин графу,  $T^F$  – наперед задана множина фреймів-прототипів;  $R$  – множина дуг  $r_k = \langle x_i, x_j \rangle$ ,  $i, j = \overline{1, n}$ ,  $r_k \in T^R$ , де  $T^R$  – наперед задана множина типів дуг. Вершини графа відповідають фреймам-екземплярам, а дуги – зв'язкам між ними.

Проте, відсутність формальної семантики ускладнює логічну інтерпретацію і порівняння властивостей різних фреймових моделей. Крім того, внаслідок типізації вершин висуваються обмеження щодо поповнення бази новими типами знань.

Застосування логічних моделей надання знань створює формальну основу теорії надання знань. Вони можуть базуватися на застосуванні сїллогістики Аристотеля, обчислення висловлювань або обчислення предикатів. Логічні моделі набули поширення на перших етапах створення СППР, що зумовлено наявністю чітких механізмів проведення логічного виводу. Проведення логічного виводу в таких моделях здійснюється на основі формальної теорії [0], яка задається четвіркою символів  $M = \langle T, P, A, B \rangle$ , де  $T$  – множина базових елементів,  $P$  – множина синтаксичних правил, за допомогою яких з елементів  $T$  створюються синтаксично правильні вирази,  $A$  – множина апріорно істинних виразів (аксіом),  $B$  – множина правил виводу нових виразів із множини  $A$ .

Логічну базу знань на основі предикатів першого порядку можна зобразити в загальному вигляді як сукупність логічних виразів виду  $P(t_1, t_2, \dots, t_n)$ , де  $t_i$  – терми предиката, які можуть бути константами, змінними або складеними термами. Кожний предикат відображає поняття

предметної області, пов'язане з сукупністю передумов, виражених термами. Можливість використовувати складені терми дає змогу відобразити знання про об'єкт прийняття рішень у вигляді логічних зв'язків між поняттями.

Переваги логічних моделей, що використовують мову логіки предикатів, пов'язані з дедуктивними можливостями обчислення предикатів: досконалий формальний апарат логічного виводу, простота реалізації процедур виявлення протиріч.

Недоліками слід вважати високі вимоги, що висуваються до предметної області, і надмірний рівень формалізації знань, що ускладнює їхнє сприйняття. Такі моделі в складних предметних областях можуть бути надто громіздкими і недостатньо наочними для аналізу знань. Знання, що містяться в базах знань, заснованих на логічних моделях, мають суто поверховий характер.

Продукційні моделі надання знань, що були вперше запропоновані Е. Постом у 70-і роки, за аналогією з логічними моделями засновуються на застосуванні формальних породжуючих граматики.

Асоціативні зв'язки між поняттями предметної області та відповідними реакціями виражаються у вигляді правил-продукцій виду „якщо – то”, що адекватно відображують процеси логічного виводу у розумі людини. Бази знань продукційного типу являють собою сукупність незалежних правил-продукцій та понять предметної області.

У загальному вигляді продукцію зображають у вигляді кортежу:

$$R = \langle I, Q, P, A \Rightarrow B, N \rangle$$

де  $I$  – ім'я продукції;  $Q$  – сфера застосування продукції (метазнання системи);  $P$  – предикат-умова застосування ядра продукції;  $A$  – умова продукції;  $B$  – наслідок продукції;  $N$  – послідовність дій після застосування продукції. Інтерпретація продукції найчастіше має вигляд „якщо – то”, а для більш складних випадків – „якщо – то – інакше”. Функція продукційної системи полягає в породженні ланцюга логічного виводу у вигляді графа. Для оптимізації процесів обробки знань можуть використовуватись процедури зворотного і змішаного виводу [3].

Продукційні моделі характеризуються такими перевагами: простота аналізу окремих продукцій, простота поповнювання бази знань, простота реалізації бази знань та висока ефективність механізмів логічного виводу для невеликих обсягів знань. У базах знань на основі продукційних моделей легко реалізується масштабованість і перевірка несуперечливості.

До недоліків продукційних моделей відносять: складність оцінки цілісного обсягу знань за окремими правилами, низька ефективність вирішення різнотипних задач; обмежена здатність продукційних систем до

навчання; низька ефективність процедур обробки знань у разі необхідності вирішення складних задач. Продукційні моделі відображають лише поверхові знання про об'єкт прийняття рішень, а для відображення глибоких, семантичних зв'язків потрібні додаткові механізми.

Нейронні мережі (рис. 2) останнім часом набули широкого поширення як метод добування та надання знань [6]. Використання нейронних мереж дає змогу опосередковано надавати знання через навчаючі вибірки та структуру мережі.

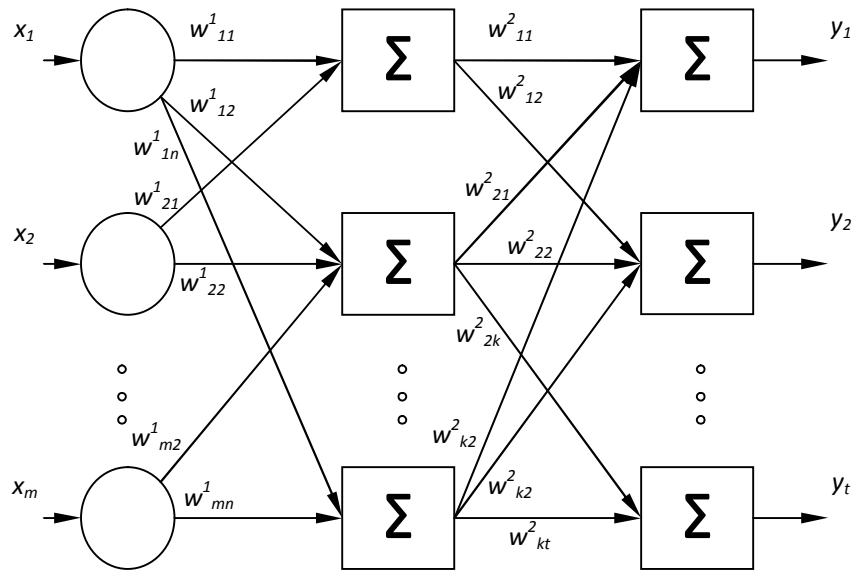


Рис. 2. Загальна структура багатошарової нейронної мережі

Нейронні мережі функціонують у двох режимах: навчання та обробки даних. У процесі навчання на вхід нейронної мережі подаються навчаючі комбінації, за якими здійснюється корекція вагових коефіцієнтів мережі. Навчання закінчується по досягненні прийняттого рівня похибок на тестових прикладах або по досягненні граничної кількості ітерацій. У процесі роботи мережі кожний нейрон підсумовує зважені вхідні сигнали та здійснює перетворення (лінійне або нелінійне) з використанням активаційної функції певного типу. Робота нейрону у мережі зображається у загальному вигляді:

$$OUT = F\left(\sum_i w_{ij}^l x_{ij}^l - \Theta\right),$$

де  $x_{ij}^l$  –  $i$ -й вхідний сигнал  $j$ -го нейрона шару  $l$ ;  $w_{ij}^l$  – ваговий коефіцієнт  $i$ -го входу  $j$ -го нейрона шару  $l$ ;  $\Theta_{jl}$  – пороговий рівень  $j$ -го нейрона шару  $l$ ;  $F$  – функція активації нейронів;  $OUT$  – вихідний сигнал нейрону.

Навчання нейронних мереж для вирішення конкретних задач полягає у виборі кількості шарів мережі, виборі виду функції активації нейронів, визначенні меж навчання мережі, а також в інтерпретації її вихідних значень.

У цілому, нейронні мережі характеризуються такими важливими позитивними ознаками, як здатність до навчання, узагальнення та абстрагування. Властивість нейронних мереж до узагальнення визначає їхню придатність до функціонування в умовах неповноти та суперечності вхідних даних, отже нейронні мережі можуть застосовуватися для вирішення задач класифікації, розпізнавання образів та суміжних з ними.

Нейронні мережі характеризуються такими негативними рисами: навчання нейронних мереж можливе лише в межах фіксованої структури вхідних даних, а сприйняття нових типів даних потребує зміни структури мережі; СППР на основі нейронних мереж не здатні пояснювати згенероване рішення; у випадку неповної навченості або невідповідності структури мережі задачам виявити помилки можна лише емпіричним шляхом.

Останнім часом активного розвитку набули квантові моделі надання знань, що базуються на принципах індуктивного моделювання. Типовими задачами, які вирішуються з використанням квантових моделей, є розпізнавання образів, прогнозування явищ, задачі добування та обробки знань. Основною особливістю квантових моделей знань є здатність навчатися за заданими сценарними прикладами та трансформуватися у квантові мережі виводу рішень із заздалегідь невідомою структурою [5].

Квантові моделі забезпечують формалізацію порцій знань у формі змістовних алгоритмічних структур даних (квантів), маніпулювання ними з використанням алгебри предикатів та векторно-матричних операторів, а також індуктивний синтез бази знань у процесі навчання за сценарними прикладами.

Квант знань першого рівня являє собою змістовну поименовану структуру даних у вигляді вектора виду:  $tk_1B = [10:0110:100]$ , де символом „:” відділені домени характеристик об’єкта прийняття рішень, а наявність „0” та „1” у компонентах домену характеризує присутність значення характеристики в певному діапазоні. Структура імовірнісних та нечітких квантів передбачає врахування відповідних характеристик знань. Кванти другого й вищих рівнів будуються на базі квантів знань першого рівня шляхом поєднання їх з використанням операторів суперпозиції.

Перевагами квантових моделей є здатність до об'єднання порцій знань з процедурами їхньої обробки, наявність потужного математичного апарату логічного виводу, здатність до ефективного навчання за сценарними прикладами. Порівняно з нейронними мережами, квантові моделі більш точно відображають генерацію людських причинно-наслідкових суджень.

До недоліків квантових моделей можна віднести складність аналізу цілісного образу знань для когнітолога за рахунок високого ступеня формалізації знань та значний обсяг роботи в процесі формування сценарних прикладів.

Одним із актуальних напрямів досліджень в області моделей надання знань є гібридизація моделей. Представником цього класу моделей є активні семантичні мережі, що запропоновані Амосовим Н.М. як засіб побудови активних систем знань [6].

З метою врахування динамічного стану об'єкта прийняття рішень на множині вузлів та зв'язків активної семантичної мережі визначено операції передачі збудження – деякої числової величини. Дана величина характеризує ступень істинності інформації в активній семантичній мережі. Стан мережі у деякий момент часу описується як стан збудження її вузлів. Зв'язки між вузлами характеризуються ваговими коефіцієнтами. Функція активації вузлів мережі представляється виразом:

$$G_i^{t+1} = (K_U - m \sum_{j=1}^n (G_j^t r_{ij}) \ln(\sum_{j=1}^n (G_j^t \tilde{r}_{ij}) + \exp(\frac{G_i^t a}{K_U})) - \Theta),$$

де  $G_i^t$  – вихідне збудження  $i$ -го вузла в момент  $t$ ;  $K_U$  – коефіцієнт збудження вузлів;  $m$  – масштабний коефіцієнт;  $r_{ij}$  – вага посилюючих вхідних зв'язків;  $\tilde{r}_{ij}$  – вага загальмовуючих вхідних зв'язків;  $a$  – коефіцієнт згасання вихідного збудження;  $\Theta$  – порогове збудження вузлів.

Перевагами даного типу моделей надання знань є можливість гнучкої адаптації бази знань до стану об'єкта прийняття рішень, висока продуктивність в поєднанні з універсальністю, а також стійкість до протиріч. На відміну від нейромережних моделей, активні семантичні мережі здатні пояснювати хід отриманих рішень.

У якості недоліків виділяється неможливість автоматичного навчання за сценарними прикладами та автоматичної формалізації даних та значну надмірність опису.

Застосування моделей надання знань при створенні СППР на сучасному етапі найчастіше базується на логіко-лінгвістичному підході [7]. Основою логіко-лінгвістичного підходу є теорія нечітких множин, яка дозволяє формалізувати нечітке вербальне описання об'єктів прийняття



рішень і застосувати відповідні математичні методи для обробки нечітких даних. Метою застосування математичного та логічного апарату теорії нечітких множин є пошук і оцінка переважних альтернатив з використанням процедур експертного опитування з врахуванням нечіткості висловлювань експертів і невизначеності вхідних даних. Вказаний підхід широко застосовується для модифікації описаних моделей надання знань.

З метою порівняння властивостей описаних моделей надання знань доцільно визначити функції підсистеми надання знань в СППР: організація зберігання знань в системі; введення нових знань і інтеграція з існуючими; виведення нових знань; пошук знань; перевірка непротиворечивості бази знань; а також здійснення інтерфейсу між користувачем і базою знань [0, 3, 7]. Для ефективного виконання наведених функцій моделі надання знань повинні задовольняти наступним вимогам: прогнозована ефективність, здатність до пояснення рішення, здатність до навчання, продуктивність, масштабованість, можливість експорту/імпорту знань, наочність.

Перша вимога визначає здатність системи оцінити якість згенерованого рішення. У разі незадовільної оцінки якості рішення система має вказати користувачу джерела суперечливості або неповноти для уточнення вхідних даних.

Здатність до пояснення згенерованого рішення є очевидною вимогою до організаційно-технічних систем управління.

Під здатністю до навчання розуміється спроможність системи в автоматизованому режимі накопичувати знання в процесі роботи та можливість враховувати нові типи вхідних даних.

Продуктивність системи – це прогнозований час її реакції на зовнішні впливи та наявність механізмів оптимізації процедур виводу.

Масштабованість моделі характеризує можливість нарощування бази знань або об'єднання декількох баз знань різних рівнів у єдину ієрархічну систему.

Можливість експорту/імпорту знань визначає придатність бази знань до модернізації в процесі еволюції системи управління.

Наочність моделі надання знань характеризує її придатність до сприйняття та аналізу когнітологами цілісного образу знань.

На основі аналізу властивостей моделей надання знань, їхніх переваг та недоліків можна зробити висновок, що кожна з них лише частково відповідає заданим вимогам .

## Висновок

Проаналізувавши методи представлення знань, можна зробити висновок, що всі існуючі моделі знань мають свої переваги та недоліки. Тому при виборі певної моделі перш за все необхідно враховувати відповідність вимогам, які ставляться перед системою для вирішення конкретної задачі.

Проаналізувавши переваги та недоліки існуючих моделей надання знань стає очевидно, що всі вони лише частково відповідають вимогам, що ставляться перед підсистемою надання знань в СППР.

Тому стає актуальною задача обґрунтованого вибору адекватної моделі надання знань у СППР. Задача за своїм змістом є багатокритеріальною та слабоформалізованою, традиційно вона вирішується на основі досвіду та кваліфікації розробників, їхніх суб'єктивних поглядів і упереджень.

Таким чином, вирішення цієї задачі є актуальною науковою проблемою і потребує подальших досліджень. Перспективним напрямком для її вирішення можна вважати застосування технології інженерії знань та теорії нечітких множин.

### *Використані джерела інформації:*

1. Трахтенгерц Э.А. Компьютерная поддержка принятия решений: Научно-практическое издание. Серия «Информатизация России на пороге XXI века» – М.: СИНТЕГ, 1988. – 376 с.
2. Norvig P. Artificial Intelligence: A Modern Approach – Prentice Hall, 2002. – 1132 p.
3. Гаврилова Т.А., Червинская К.Р. Извлечение и структурирование знаний для экспертных систем. – М.: Радио и связь, 1992. – 200 с.
4. Нейросетевые алгоритмы прогнозирования и оптимизации систем / Назаров А.В., Лоскутов А.И. (Под ред. Финкова М.В.) – СПб: Наука и техника, 2003. – 384 с.
5. Сироджа И.Б. Квантовые модели и методы искусственного интеллекта для принятия решений и управления – К.: Наукова думка, 2002. – 490 с.
6. Нейрокомпьютеры и интеллектуальные роботы / Амосов Н.М., Байдык Т.Н., Гольцев А.Д. и др. (Под ред. Амосова Н.М.): АН УССР Институт кибернетики – К.: Наукова думка, 1991. – 272 с.
7. Герасимов Б.М., Тарасов В.А., Токарев И.В. Человеко-машинные системы принятия решений с элементами искусственного интеллекта – К.: Наукова думка, 1993. – 184 с.

*Рецензент: д.т.н. Крижний А.В.*