

УДК 81'33:004.81

Бегаль О. В.

## КОГНІТИВНЕ МОДЕЛЮВАННЯ ПАМ'ЯТІ АВТОМАТИЧНИХ ДІАЛОГОВИХ СИСТЕМ

*Статтю присвячено дослідженню когнітивних аспектів моделювання пам'яті автоматичних діалогових систем із метою створення віртуальних співрозмовників, здатних успішно підтримати діалог із людиною. Запропоновано підхід до структуривання пам'яті лінгвістично компетентного чат-бота та схему результативного пошуку даних у базі знань діалогової системи.*

**Ключові слова:** когнітивне моделювання пам'яті, моделі пам'яті, віртуальний співрозмовник, комунікативна взаємодія, діалог.

**Бегаль Е. В. Когнитивное моделирование памяти автоматических диалоговых систем. – Статья.**

*Статья посвящена исследованию когнитивных аспектов моделирования памяти автоматических диалоговых систем с целью создания виртуальных собеседников, способных успешно поддержать диалог с человеком. Предложены подход к структурированию памяти лингвистически компетентного чат-бота и схема результативного поиска данных в базе знаний диалоговой системы.*

**Ключевые слова:** когнитивное моделирование памяти, модели памяти, виртуальный собеседник, коммуникативное взаимодействие, диалог.

**Begal O. V. Cognitive modeling of the memory for automatic dialogue systems. – Article.**

*This article is devoted to cognitive aspects of modeling the memory for the automatic dialogue systems in view of developing virtual interlocutors able to maintain a dialogue with humans. The approach to structuring the memory of linguistically competent chat-bot is suggested, as well as the scheme for effective data search in the knowledge base of the dialogue system.*

**Key words:** cognitive memory modeling, models of memory, virtual interlocutor, communicative interaction, dialogue.

Успішне виконання комп'ютером комунікативних і когнітивних функцій, притаманних людині, є одним із ключових завдань сучасних досліджень у галузі штучного інтелекту. З метою поглибленого й вичерпного аналізу когнітивних властивостей комунікативної взаємодії слід звернути увагу на моделі пам'яті, які у свою чергу дають змогу виокремити аспекти, важливі для роботи віртуального співрозмовника. Обрання ефективної й дієвої моделі пам'яті для розробки алгоритмів зберігання та впорядкування даних із їх подальшим виведенням і застосуванням є запорукою результативного оперування машини в діалозі з людиною.

Науковцями було запропоновано різноманітні підходи до моделювання пам'яті, які трактують ключові аспекти й параметри її впорядкування, а також механізми реалізації пошуку даних. Зокрема, К. Коффка [3] сформулював неасоціативну теорію людської довгострокової пам'яті, яку піддав критиці В. Вікельгреном [9]. Загальноприйнятою вважається асоціативна мережева модель, що стала поштовхом до оптимізації роботи машин.

Було запропоновано теорії пошуку даних у пам'яті, серед яких на особливу увагу заслуговують теорія ієрархічного пошуку, обґрунтована В. Вікельгреном, і теорія поширюваної активації семантичної обробки даних, розроблена М.Р. Квілліаном [6; 7; 8] та розвинена А. Коллінзом та Е. Лофтусом [2].

Об'єкт дослідження – діалогічна взаємодія людини та комп'ютера (автоматичної діалогової системи).

Предмет дослідження – когнітивні аспекти моделювання пам'яті віртуального співрозмовника.

Мета статті полягає у встановленні принципів ефективного моделювання пам'яті автоматичної

діалогової системи та розробці схеми результативного пошуку даних у базі знань віртуального співрозмовника.

Завданнями дослідження є характеристика ключових підходів до моделювання пам'яті й реалізації пошуку в її межах, а також формування підходу до ефективного структуривання пам'яті та схеми результативного пошуку даних у базі знань автоматичної діалогової системи.

Розглянемо докладніше основні підходи до моделювання пам'яті. К. Коффка запропонував неасоціативну теорію людської довгострокової пам'яті, згідно з якою розум перебуває в постійному процесі кодування й збереження повсякденного досвіду, формуючи при цьому так звані «стовпець залишених у пам'яті слідів» (*trace column*) [3]. Виходячи з наведеного трактування структури пам'яті, вона являє собою нагромадження концептів у лінійному порядку, які продовжують надходити в процесі отримання нових знань. Це означає, що сукупність концептів, які містяться в пам'яті, не пов'язані між собою асоціативними ланками, а лише представляють сприйнятту ситуацію. Цей підхід до пояснення особливостей влаштування пам'яті критикує В. Вікельгреном, зауважуючи, що одним із найбільших здобутків когнітивної психології є спростування неасоціативної теорії довгострокової пам'яті [9, 22].

З-поміж існуючих моделей пам'яті загально-визнаною вважається *асоціативна мережева модель* (*associative network model*) [5]. Згідно із цією моделлю між концептами існують зв'язки, які різняться за рівнем асоціативних сполучень. Це означає, що деякі концепти формують більш стійкі асоціативні ряди, ніж інші. У нашій сві-

домості знання структуровані таким чином, що асоціативні ланки, що існують між концептами, дозволяють переходити від одного концепту до іншого на основі смислових зв'язків між ними. Увійшовши в певну концептуальну область, ми отримуємо доступ до всіх наявних у ній даних. Шляхом переміщення асоціативними рядами ми оперуємо концептами, поєднуємо їх, встановлюємо співвідношення, виводимо логічні висновки та формуємо нові судження.

У межах дослідження ми здійснюємо розробку лінгвістично компетентного віртуального співрозмовника, здатного в режимі природного діалогу з людиною забезпечити її необхідною інформацією про вченого певного періоду та напряму мовознавства. Автоматична діалогова система охопить широкий спектр знань про лінгвіста, що включає такі складові компоненти образу вченого: його життя (день і місце народження/смерті, сім'я, освіта, статус, професія); ключові концепції, теорії та ідеї; наукові праці. Розглянемо особливості асоціативної мережевої моделі пам'яті на прикладі схеми, на якій відтворено структуру знань про сім'ю Н. Хомського (див. схему 1).

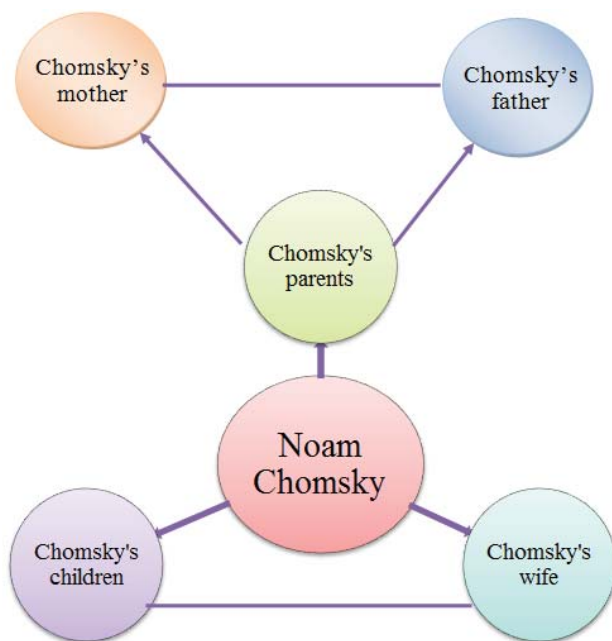


Схема 1.  
Асоціативна мережева модель пам'яті

Центральну позицію в цій схемі займає концепт, який представляє постать Н. Хомського. Якщо задіяти концептуальну область сімейних стосунків, то в асоціативному відношенні постануть концепти його батьків, дружини й дітей. У безпосередньому зв'язку з концептом БАТЬКИ перебувають концепти МАТИ та БАТЬКО, які також пов'язані між собою. Окрім цього, активація концепту ДРУЖИНА зумовлює відповідне посилення на концепт ДІТИ. У поданій схемі відобра-

жено мережеву єдність знань, які вибудовують цілий асоціативний ряд під час фокусування на темі родини Н. Хомського.

У галузі штучного інтелекту асоціативна пам'ять є особливим видом пам'яті, що адресується за вмістом (*content-addressable memory* або *CAM*) та наділена властивістю прямого доступу до даних. На відміну від асоціативної пам'яті довільного доступу (*Random Access Memory* або *RAM*) адресується за розміщенням, тобто машина здійснює пошук за всіма адресами пам'яті, щоб знайти те місце, де зберігаються дані, які відповідають запиту. Асоціативна пам'ять влаштована так, що машина прямо працює зі змістом масиву даних, а не з адресами їх зберігання [9, 23].

На думку В. Вікельгрена, найбільш доцільною альтернативою прямому доступу до змістового наповнення є *теорія ієрархічного пошуку* (*hierarchical search theory*), згідно з якою спершу пошук здійснюється з-поміж категорій найвищого рівня, і коли необхідну буде знайдено, то в її межах із категорій нижчого рангу обиратиметься та, що відповідає запиту, і так далі, доки потрібні дані не буде віднайдені [9, 23].

У деяких теоріях семантичної пам'яті створюється новий структурний вузол (*node*) для кожної появи відповідного концепту в нових судженнях, коли формуються зв'язки з іншими концептами. Закодовуючи кожен набір концептів, який складає окреме судження, ми отримуємо можливість уникнути проблем, пов'язаних із накладанням асоціативних ланок та інтерференцією. В. Вікельгрен наводить такий приклад: якщо закодувати вислови «*John hit Frank*» і «*Peter hit Bill*» як асоціативні ланки для того самого вузла «*hit*», то в результаті ми отримаємо пам'ять, яка не зможе остаточно вивести: Джон вдарив Френка чи Білла [9, 26].

Спробу відтворити механізм пошуку даних у пам'яті з точки зору комп'ютера здійснив М.Р. Квілліан [6; 8]. Він запропонував *теорію поширюваної активації семантичної обробки даних* (*a spreading-activation theory of semantic processing*), у межах якої пошук у пам'яті розглядається як активація, що поширюється від двох або більше вузлів концептів у семантичній мережі, доки не буде віднайдені місце їх перетину. Зазначену теорію згодом переглянули й доповнили А. Коллінз та Е. Лофтус [2].

Концепт може бути представлено як окремий вузол у мережі, котрий з'єднаний ланками з іншими концептами. Ланки можуть мати різний рівень семантичної та асоціативної релевантності, а також когнітивної ваги (*criticalities*), котрий вказує, наскільки суттєвою й важливою є кожна ланка для значення концепту [2, 408].

М.Р. Квілліан розмежовує п'ять основних типів зв'язків між концептами [7]:

- 1) надкласові/підкласові (*ISA relationship*);
- 2) модифікаційні (наприклад, прикметник чи прислівник);
- 3) розділові (логічний оператор «*OR*»);
- 4) єднальні (логічний оператор «*AND*»);
- 5) клас зв'язків, що представляє ті зв'язки (зазвичай дієслівні), які являють собою окремий концепт.

Пошук у пам'яті передбачає паралельне відстежування ланок від вузла кожного з концептів та зумовлений вхідними словами-стимулами. Зона активації поступово розширюється, спочатку до всіх вузлів, які з'єднані з першим вузлом, далі – до всіх наступних, які мають зв'язок із кожним із цих вузлів.

Кожному вузлу, до якого дійде пошук, приписується активаційний тег, яким відзначається початковий вузол і попередній. Коли цей тег перетнеться з тегом іншого початкового вузла, пошук буде завершено. Наявність тегів дозволяє реконструювати весь шлях, який привів до перетину. Після того, як перетин було знайдено, необхідно оцінити, чи задовольняє отриманий результат усі параметри й вимоги синтаксису та контексту вхідного запиту [2, 408].

Чим більше спільних властивостей і якостей мають концепти між собою, тим більше зв'язків утворюється між їхніми вузлами, а отже, це безпосередньо впливатиме на близькість семантичної взаємодії цих концептів [2, 411]. Рівень міцності зв'язку залежатиме від кількості спільних смислових елементів і характеристик. А. Коллінз та Е. Лофтус ілюструють особливості реалізації описаного механізму встановлення взаємозв'язків між концептами за допомогою схеми (див. схему 2) [2, 412].

У поданій схемі відображено структуру взаємозв'язків між концептами, де коротші лінії сигналізують про вищий рівень семантичної взаємодії. Смислова віддаленість свідчить про слабкий асоціативний зв'язок. Наприклад, якщо на перший план вивести зв'язок концептів «*RED*» та «*FIRE ENGINE*», то активація не поширяться на концепти «*CHERRIES*», «*ROSES*» чи «*SUNSETS*», що зумовлено малою кількістю смислових ланок, які єднали би їх із концептом «*FIRE ENGINE*». З останнім в асоціативному ряді взаємодіятимуть такі концепти, як «*BUS*», «*TRUCK*» або «*AMBULANCE*».

Проаналізувавши особливості механізму пошуку даних у пам'яті, наведеного в межах теорії поширеної активації семантичної обробки, ми можемо зробити висновок, що вона являє собою зручну модель для відтворення ментальних процесів в оперуванні комп'ютера в ролі віртуального співрозмовника. Виокремлені суттєві складники й важливі аспекти зв'язку між ними дадуть змогу вибудувати дієву стратегію розробки чат-бота, здатного до комунікативної діяльності.

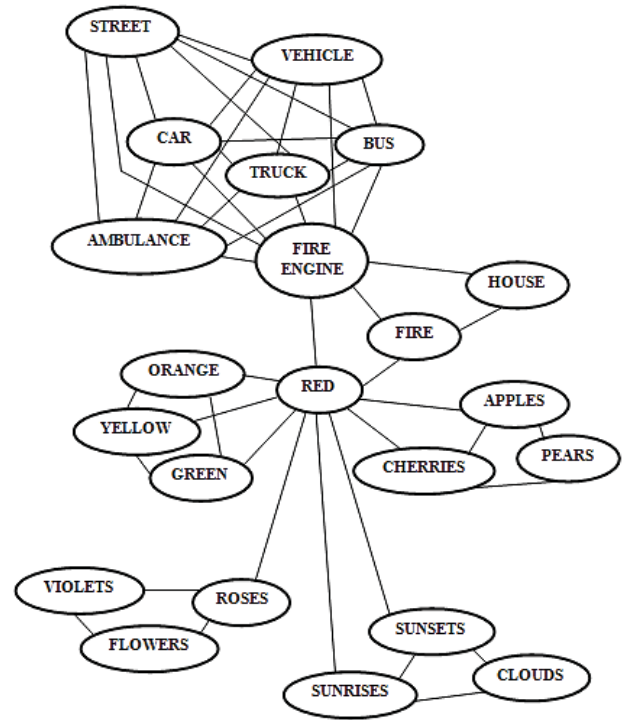


Схема 2. Схематична репрезентація семантичної взаємодії концептів

Беручи до уваги розглянуті підходи до моделювання пам'яті, ми виокремили ключові параметри, які необхідно врахувати в розробці автоматичної діалогової системи, здатної успішно виконувати роль активного учасника комунікативної взаємодії, а саме:

1. *Асоціативна мережа* апелює до асоціативної мережевої моделі пам'яті, що є ключовим підходом до структурування й збереження даних у базі знань, дасть змогу впорядкувати знання з певної предметної області у вигляді чіткої фреймової ієрархічної організації із вкладеними слотами (тобто вкладеними даними). Наприклад, для фрейму сім'ї вченого слотами будуть конкретні дані-репліки про його сім'ю. Фрейми мають бути поєднані зв'язками, що надасть можливість переходити від загальної інформації до конкретної, при цьому необхідно запрограмувати систему на «запам'ятовування» шляху просування від теми до підтеми. Фрейм розглядається як модель структурованої репрезентації знань (за термінологією М. Мінського [4]), що за своєю будовою формує деревовидну ієрархію (в інформатиці поширеною є назва «онтологія» [1]). Фрейм міститиме інформацію, яка розгортатиметься на декількох рівнях: від найбільш загального до конкретного.

2. *Загальний доступ до даних*: увійшовши до певної концептуальної області через асоціативні зв'язки, чат-бот отримуватиме доступ до всіх даних, які містяться в ній.



3. *Створення структурних вузлів* маркування нового структурного вузла для кожної появи відповідного концепту, що дасть змогу уникнути накладання асоціативних ланок.

4. *Збереження історії діалогу*: необхідно зберегти історію розгортання діалогічної комунікації для орієнтації віртуального співрозмовника в контексті розмови, вдалого встановлення референційних зв'язків та дотримання семантичної й прагматичної єдності діалогу.

5. *Ієрархічний пошук*: услід за теорією ієрархічного пошуку спочатку віртуальний співрозмовник здійснюватиме пошук серед категорій найвищого рівня, поступово просуваючись до нижчих рівнів. Рух від загального до конкретного є запорукою входження в потрібну фреймову область знань.

6. *Адресація за вмістом* забезпечить прямий доступ до даних.

7. *Пошук найбільшої концентрації слів із найвищим когнітивним статусом*: ключовим для пошуку даних у пам'яті віртуального співрозмовника є видача репліки, яка найбільшою мірою асоціюється з ключовими поняттями запиту. Для забезпечення відповідності реакції стимулу необхідно віднайти той слот бази знань, де ключові слова з найвищим когнітивним статусом і вагою зосереджені в найбільшій кількості, тобто встановити найвдаліше місце їх перетину.

Керуючись сформованим підходом до когнітивного моделювання пам'яті, ми здійснили розробку лінгвістично компетентного віртуального співрозмовника, здатного підтримати розмову про лінгвістів різних періодів і напрямів мовознавства, їх праці та основні концепції. Усі дані репрезентовані у вигляді онтологічного представлення знань, що утворює деревовидну ієрархію. Розглянемо приклад моделі знань для Н. Хомського (див. схему 3).

У поданій схемі відтворено структуру впорядкування знань про Н. Хомського в базі даних автоматичної діалогової системи. Деревовидна ієрархічна організація відповідає ключовим принципам асоціативної мережевої моделі пам'яті, забезпечує ефективний пошук релевантних реакцій на запропонований людиною стимул.

Експериментальним шляхом підтверджено, що результативність пошуку досягається не під час добору найбільшої кількості збігів між реплікою-стимулом і можливими репліками-реакціями, а коли задіюється схема аналізу вхідної репліки користувача, скерована на відбір даних за найменшою кількістю розбіжностей запиту з прототипом питань, що можуть задіяти відповідну репліку-реакцію, і навпаки – прототипу питань із вхідним запитом. Обрахунок здійснюється за такою формулою:

$$diff=(1-koef1)+(1-koef2),$$

де *koef1* – коефіцієнт збігів запиту з прототипом питань, *koef2* – коефіцієнт збігів прототипу питань із вхідним запитом,  $(1-koef1)$  – коефіцієнт розбіжностей запиту з прототипом питань,  $(1-koef2)$  – коефіцієнт розбіжностей прототипу питань із вхідним запитом.

Ми прагнемо отримати якомога меншу суму розбіжностей. Це дає змогу формалізувати процес пошуку, при цьому застосовуються прийоми комп'ютерної обробки природної мови з метою підвищення якості аналізу змістового наповнення вхідних реплік людини.

Розглянемо приклад діалогу з лінгвістично компетентним віртуальним співрозмовником, у якому втілено наведені принципи когнітивного моделювання пам'яті:

*Bot: Hi, I'm Joseph Linguiston! Enter your name, please.*

*I: My name is Helen.*

*Bot: Pleasure to meet you, Helen. What linguist would you like to know about?*

*I: Let's talk about Noam Chomsky.*

*Bot: Feel free to ask me about Noam Chomsky's: life, notable ideas, main concepts, theories and works.*

*I: What are his main concepts?*

*Bot: Chomsky's main achievements are: transformational grammar theory, generative grammar, context-free grammar, the universal grammar theory, the Chomsky hierarchy and his theories on language acquisition [http://en.wikipedia.org/wiki/Noam\_Chomsky].*

*I: Tell me about transformational grammar.*

*Bot: Transformational grammar is a system of language analysis that recognizes the relationship among the various elements of a sentence and among the possible sentences of a language and uses processes or rules (some of which are called transformations) to express these relationships [http://www.britannica.com/EBchecked/topic/602647/transformational-grammar]. Would*

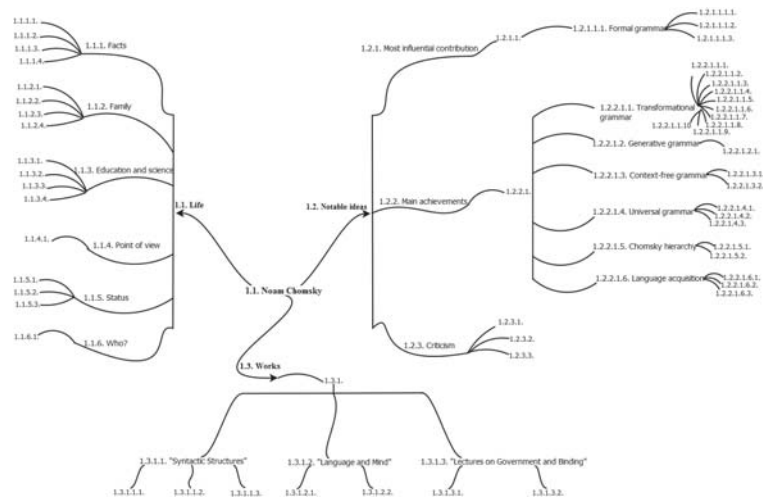


Схема 3. Модель знань про постать Н. Чомського

*you like to know more about transformational grammar?*

*I: Yes, please.*

*Bot: The main concepts of transformational grammar are a Deep Structure and a Surface Structure, I-Language and E-Language, grammaticality and a transformation [http://en.wikipedia.org/wiki/Transformational\_grammar].*

*I: Go on.*

*Bot: The Deep Structure represents the core semantic relations of a sentence, and is mapped on to the Surface Structure via transformations. For example, the sentences "Pat loves Chris" and "Chris is loved by Pat" mean roughly the same thing and use similar words. These two sentences are distinct surface forms that derive from a common Deep Structure [http://en.wikipedia.org/wiki/Transformational\_grammar].*

*I: And what is Surface structure?*

*Bot: The Surface Structure is the scope of sentences that are actually uttered [http://en.wikipedia.org/wiki/Transformational\_grammar].*

*I: What is the difference between the Deep Structure and the Surface Structure?*

*Bot: The Deep Structure represents fundamental relationships among linguistic components, while the Surface Structure is the sentences that are actually uttered. One Deep Structure can stand for many Surface Structures [http://en.wikipedia.org/wiki/Transformational\_grammar].*

*I: Thank you.*

*Bot: I hope it was a fruitful conversation.*

*I: Bye!*

*Bot: I was glad to have a talk with you, Helen.*

Такий діалог можна вважати успішним як з огляду на досягнення людиною ілюкотивної мети

комунікативного акту (отримати бажану інформацію про Н. Хомського), так і з точки зору дотримання прагматичної й семантичної зв'язності, адже віртуальний співрозмовник доцільно відповідає на поставлені запитання, дотримується теми розмови та орієнтується в цій предметній галузі.

У межах дослідження ми розглянули основні підходи до моделювання пам'яті. На основі отриманих результатів сформулювали підхід до структурування пам'яті автоматичної діалогової системи та схему результативного пошуку даних на прикладі розробки лінгвістично компетентного віртуального співрозмовника (під лінгвістичною компетентністю розуміється здатність успішно підтримати розмову про відомих лінгвістів, їх життя й основні наукові концепції).

Результати проведеного дослідження дають підстави зробити висновок, що моделювання пам'яті автоматичної діалогової системи на основі асоціативної мережевої моделі та принципу ієрархічного пошуку даних забезпечує ефективне виконання машиною ролі співрозмовника. Важливо також врахувати місце перетину ключових слів із найвищим когнітивним статусом і вагою, що дає змогу реалізувати успішний діалог між людиною та комп'ютером. Сформований підхід до моделювання пам'яті чат-бота та схема пошуку даних у базі знань успішно втілюються в розробці лінгвістично компетентних віртуальних співрозмовників.

Перспективи подальших досліджень полягають у тому, що розвиток і впровадження асоціативно-мережевих підходів до моделювання пам'яті дасть змогу оптимізувати роботу автоматичних діалогових систем для результативної комунікації з людиною.

### Література

1. Лапшин В.А. Онтология в компьютерных системах / В.А. Лапшин. – М. : Научный мир, 2010. – 224 с.
2. Collins A.M. A spreading-activation theory of semantic processing / A.M. Collins, E.F. Loftus // *Psychological Review*. – 1975. – Vol. 82. – № 6. – P. 407–428.
3. Koffka K. Principles of Gestalt Psychology / K. Koffka. – New York : Harcourt, 1935. – 485 p.
4. Minsky M. A Framework for Representing Knowledge (A.I. Memo 306) / M. Minsky. – Cambridge, MA : Artificial Intelligence Laboratory ; Massachusetts Institute of Technology, 1974. – 81 p.
5. Mitchell A.A. Models of Memory: Implications for Measuring Knowledge Structures / A.A. Mitchell, A. Abor (Eds.) // *Advances in Consumer Research*. – MI: Association for Consumer Research. – 1982. – Vol. 9. – P. 45–51.
6. Quillian M.R. A revised design for an understanding machine / M.R. Quillian // *Mechanical Translation*. – 1962. – Vol. 7. – P. 17–29.
7. Quillian M.R. Semantic memory / M.R. Quillian, M.L. Minsky (Ed.) // *Semantic information processing*. – Cambridge, MA : The MIT Press, 1968. – P. 227–270.
8. Quillian M.R. Word concepts: A theory and simulation of some basic semantic capabilities / M.R. Quillian // *Behavioral Science*. – 1967. – Vol. 12. – P. 410–430.
9. Wickelgren W.A. Human learning and memory / W.A. Wickelgren // *Annual Review of Psychology*. – 1981. – Vol. 32. – P. 21–52. – [Electronic resource]. – Mode of access : <http://www.columbia.edu/~nvg1/Wickelgren/papers/1981aWAW.pdf>.
10. Encyclopædia Britannica Online [Electronic resource]. – Mode of access : <http://www.britannica.com/EBchecked/topic/602647/transformational-grammar>.
11. Wikipedia website [Electronic resource]. – Mode of access : <http://www.wikipedia.org>.