

УДК 510.635:004.891(045)

doi:10.20998/2413-4295.2016.12.14

СПОСОБИ ВИЯВЛЕННЯ ЛОГІЧНИХ ЗВ'ЯЗКІВ МІЖ ЧАСТИНАМИ ТЕКСТОВИХ ДОКУМЕНТІВ

А. І. ВАВІЛЕНКОВА

*Кафедра комп'ютеризованих систем управління, Національний авіаційний університет, Київ, УКРАЇНА
e-mail: a_vavilenkova@mail.ru*

АНОТАЦІЯ Метою роботи є створення формального опису способів виявлення логічних зв'язків між частинами текстів. Одним із засобів лінгвістичного аналізу текстової інформації та екстракції знань є логіко-лінгвістичні моделі. Довільний текст можна представити у вигляді масиву логіко-лінгвістичних моделей речень природної мови, логічні операції та структурні компоненти яких вказують на типи логічних зв'язків. Пропонується для відображення структури тексту використовувати формальний опис п'яти абстрактних моделей, що в подальшому дають можливість визначити тематику тексту.

Ключові слова: логіко-лінгвістичні моделі, природна мова, логічні зв'язки, текстовий документ, аналіз.

METHODS OF IDENTIFYING LOGICAL CONNECTIONS BETWEEN PARTS OF TEXT DOCUMENTS

A. VAVILENKOVA

Department of Computerized Control Systems, National Aviation University, Kyiv, UKRAINE

ABSTRACT The objection of work is creation a formal specification of methods of identifying logical connections between parts of texts. One of the means of text information linguistic analysis and knowledge extraction is logic and linguistic model. Any text can be presented as array of logic and linguistic models of natural language sentences, where logic operations and structural components indicate the types of logical relationships. There are two levels while constructing the text. The first level reflects the grammar links between the natural language sentences, and the second one is in content relationships between fragments of text. Exactly the construction of the second level is proposed to be used for formal description of five abstract models. For that purpose, every natural language sentence transforms into logic and linguistic model, which is an atomic predicate for the simple sentence and, correspondingly, set of atomic predicates for the complex sentence. The first abstract model is based on the consistent deployment of information when the object of the preceding sentence becomes subject for the next sentence. For the second abstract model the key point is usage of the same subject in a few natural language sentences, and within these sentences subject is represented in different interpretations. The third model is based on the fact that the subjects of logic and linguistic models of the next natural language sentences are hyperons or conversives in relation to the subject of the first sentence of text fragment. For the fourth abstract model the subject of the first sentence is partly the subject for the next sentences. In the fifth model the subjects of the next sentences can be both objects and matters of the first sentence and its subject. The result of the construction of such abstract models are polyhedrons that subsequently used to determine the texts' subject (topic).

Keywords: logic and linguistic models, natural language, logical links, text document analysis.

Вступ

Збільшення інформаційних потоків та динамічний характер інформаційного попиту спричиняє потребу у постійному отриманні оперативної інформації для своєчасного прийняття правильних рішень [1]. Тому першочерговою задачею при обробці великих обсягів інформації є аналіз даних, а системи управління базами даних повинні підтримувати методи повнотекстового пошуку, які б дозволили швидко знаходити потрібну інформацію серед великих об'ємів текстів [2]. Аналіз даних передбачає створення математичних методів та алгоритмів екстракції знань з експериментальних даних; фільтрацію, перетворення та моделювання даних з метою отримання корисної інформації та прийняття рішень [3].

Дослідження текстових документів на змістовному рівні передбачає застосування методів

лінгвістичного аналізу та основних положень теорії тексту [4]. Існуючі системи лінгвістичного аналізу (Oracle Text, Advego, IBM Intelligent Miner for Text, технологія ABBY Y Compreno, Cognitive Dwarf та ін.) оперують різноманітними парсерами, механізмами перекладу, автоматичного реферування та анотування, тезаурусами [5]. Проте жоден із перерахованих програмних продуктів не здійснює змістовного аналізу та автоматичного пошуку логічних зв'язків між складними синтаксичними конструкціями текстів. У свою чергу, фахівці у сфері структурної лінгвістики (В.А. Широков [6], Дж. Лайонз [7], І.Р. Гальперін [8], К.А. Філіппов [9]) виділяють та описують форми змістовного зв'язку у текстах, не підкріплюючи їх алгоритмами автоматичного простеження. Тому актуально є задача створення формального опису способів організації логічних зв'язків у текстах для подальшого автоматичного аналізу текстових документів.

Одним із засобів лінгвістичного аналізу текстової інформації та екстракції знань є логіко-лінгвістичні моделі. Зокрема, представлення тексту у вигляді масиву логіко-лінгвістичних моделей речень природної мови та встановлення закономірностей між їх структурними компонентами дає можливість простежити логічні зв'язки між частинами текстових документів.

Мета роботи

Метою роботи є створення формального опису способів виявлення логічних зв'язків між частинами тексів.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

- 1) Дослідити можливі способи побудови текстів довільної тематики.
- 2) Представити речення природної мови за допомогою логіко-лінгвістичних моделей.
- 3) Сформувати абстрактні моделі побудови логічних зв'язків у текстах.

Викладення основного матеріалу

Розрізняють два рівні побудови тексту. На першому рівні речення тексту зв'язуються за допомогою граматичних залежностей. Другий рівень необхідний для формування змістовної єдності тексту, що вбачається у змістовних зв'язках між фрагментами тексту, обумовлюючи його тип. Таким чином, кожен текст складається зі змістовних одиниць, які виконують у тексті певні функції.

За Т. ван Дейком модель пояснення процесів, пов'язаних зі сприйняттям змісту тексту, відбувається через послідовний зв'язок макроструктур (рис. 1).

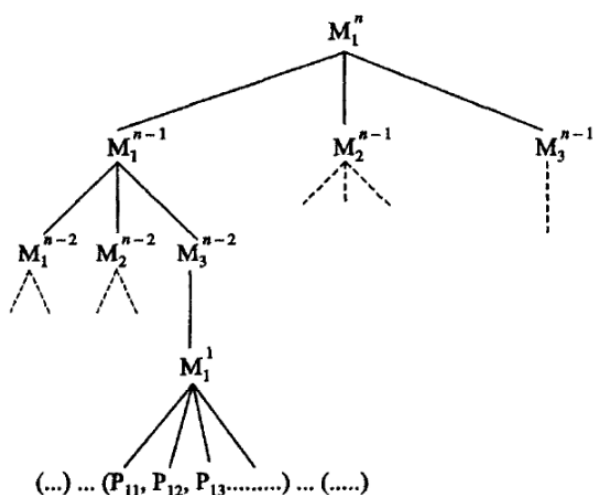


Рис. 1 – Ієрархія макроструктур в концепції Т. ван Дейка

Кожен рівень позначається більш високим індексом, охоплює ряд пропозицій до найвищого

рівня (від $M^{(1)}$ до $M^{(n)}$). Основною проблемою побудови структури тексту з використанням такої моделі є шляхи виявлення макроструктур у тексті. Це стає можливим завдяки застосуванню правил семантичної трансформації.

В основі змістовної єдності текстів довільної тематики та структури, тобто при формування другого рівня, лежить п'ять абстрактних моделей, кожна з яких представляє собою сукупність пресупозицій, компонент змісту речень, які повинні бути істинними для того, щоб речення не сприймалося як аномальне або недоречне у даному контексті [9]. Для кожної з абстрактних моделей пропонується формалізований опис способів організації зв'язків між макроструктурами тексту.

Довільне просте речення природної мови можна представити у вигляді атомарного предикату виду:

$$L(S) = p(x, g, y, q, z, r, h), \quad (1)$$

- де x – суб'єкт інваріантної до $L(S)$ логіко-лінгвістичної моделі;
- g – характеристика суб'єкта x ;
- y – об'єкт логіко-лінгвістичної моделі;
- q – характеристика об'єкта y ;
- p – відношення, що пов'язує суб'єкт x з об'єктом у логіко-лінгвістичної моделі;
- z – предмет p -го відношення між суб'єктом x та об'єктом у логіко-лінгвістичної моделі;
- r – характеристика предмету відношення;
- h – характеристик відношення p .

Істинність логічного висловлювання (1) нерозривно пов'язана з правильністю побудови речення природної мови. Зміст різних частин складного твердження (сукупності атомарних предикатів виду (1), пов'язаних логічними операціями) задається всіма можливими способами рекурсивного задання групи функцій інтерпретації, що проєктують речення природної мови на деякі завчасно задані математичні множини. Так, інтерпретація логіки предикатів першого порядку задається проєктуванням в терми, а проєктування предикатів – у значення «істина» або «хибність».

Абстрактна модель №1. Базується на послідовному розгортанні інформації, коли об'єкт або предмет попереднього речення стає суб'єктом для наступного. У такому тексті розгортання змісту відбувається від заданого до нового. Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$L(S^{(1)}) = p^{(1)}(x^{(1)}, g^{(1)}, y^{(1)}, q^{(1)}, z^{(1)}, r^{(1)}, h^{(1)}) \&$$

$$L(S^{(2)}) = p^{(2)}(y^{(1)}, g^{(2)}, y^{(2)}, q^{(2)}, z^{(2)}, r^{(2)}, h^{(2)}) \&$$

.....

$$L(S^{(i)}) = p^{(i)}(y^{(i-1)}, g^{(i)}, y^{(i)}, q^{(i)}, z^{(i)}, r^{(i)}, h^{(i)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(n)}) = p^{(n)}(y^{(n-1)}, g^{(n)}, y^{(n)}, q^{(n)}, z^{(n)}, r^{(n)}, h^{(n)}),$$

де $i = \overline{1, n}$ – номер речення у фрагментів тексту,
 n – загальна кількість речень у фрагменті тексту.

Графічна інтерпретація першої абстрактної моделі побудови тексту для фрагменту із трьох простих речень зображена на рис.2.

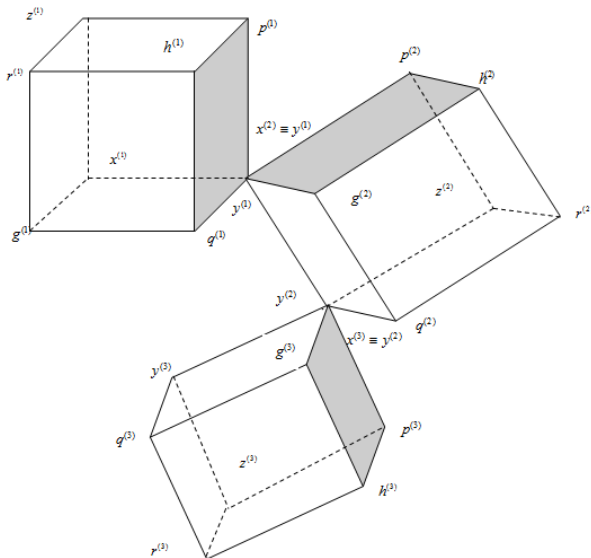


Рис. 2 – Інтерпретація першої абстрактної моделі

Абстрактна модель №2. Друга абстрактна модель характеризується тим, що у декількох реченнях природної мови вживається один і той самий суб'єкт, представлений у різних інтерпретаціях. Тобто одна і та сама тема пронизує увесь текст.

Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$L(S^{(1)}) = p^{(1)}(x^{(1)}, g^{(1)}, y^{(1)}, q^{(1)}, z^{(1)}, r^{(1)}, h^{(1)}) \&$$

$$L(S^{(2)}) = p^{(2)}(x^{(1)}, g^{(2)}, y^{(2)}, q^{(2)}, z^{(2)}, r^{(2)}, h^{(2)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(i)}) = p^{(i)}(x^{(1)}, g^{(i)}, y^{(i)}, q^{(i)}, z^{(i)}, r^{(i)}, h^{(i)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(n)}) = p^{(n)}(x^{(1)}, g^{(n)}, y^{(n)}, q^{(n)}, z^{(n)}, r^{(n)}, h^{(n)}).$$

Абстрактна модель №3. Для третьої моделі характерно, що суб'єкти логіко-лінгвістичних моделей всіх наступних речень природної мови є гіперонімами або конверсивами до суб'єкта, об'єкта, предмета відношень або характеристики відношення першого речення фрагменту, що розглядається. Таким чином, у кожній логіко-лінгвістичній моделі присутні елементи загальної тематичної направленості тексту. Як і для перших двох типів абстрактних моделей,

логіко-лінгвістичні моделі природної мови будуть з'єднані логічною операцією кон'юнкції і матимуть наступний вигляд:

$$L(S^{(1)}) = p^{(1)}(x^{(1)}, g^{(1)}, y^{(1)}, q^{(1)}, z^{(1)}, r^{(1)}, h^{(1)}) \&$$

$$L(S^{(2)}) = p^{(2)}(\tilde{x}^{(1)}, g^{(2)}, y^{(2)}, q^{(2)}, z^{(2)}, r^{(2)}, h^{(2)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(i)}) = p^{(i)}(\hat{x}^{(1)}, g^{(i)}, y^{(i)}, q^{(i)}, z^{(i)}, r^{(i)}, h^{(i)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(n)}) = p^{(n)}(\tilde{x}^{(1)}, g^{(n)}, y^{(n)}, q^{(n)}, z^{(n)}, r^{(n)}, h^{(n)}),$$

якщо $\tilde{x}^{(1)}$ – синонім суб'єкта $x^{(1)}$, $\hat{x}^{(1)}$ – гіперонім суб'єкта $x^{(1)}$, $\tilde{x}^{(1)}$ – конверсив суб'єкта $x^{(1)}$.

Якщо загальна тематика фрагменту тексту стосується предмета відношень першого речення, тобто $\hat{z}^{(1)}$ – гіперонім предмета відношень $z^{(1)}$, $\tilde{z}^{(1)}$ – синонім предмета відношень $z^{(1)}$, $\tilde{z}^{(1)}$ – конверсив предмета відношень $z^{(1)}$, то логіко-лінгвістичні моделі матимуть вигляд:

$$L(S^{(1)}) = p^{(1)}(x^{(1)}, g^{(1)}, y^{(1)}, q^{(1)}, z^{(1)}, r^{(1)}, h^{(1)}) \&$$

$$L(S^{(2)}) = p^{(2)}(\hat{z}^{(1)}, g^{(2)}, y^{(2)}, q^{(2)}, z^{(2)}, r^{(2)}, h^{(2)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(i)}) = p^{(i)}(\tilde{z}^{(1)}, g^{(i)}, y^{(i)}, q^{(i)}, z^{(i)}, r^{(i)}, h^{(i)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(n)}) = p^{(n)}(\tilde{z}^{(1)}, g^{(n)}, y^{(n)}, q^{(n)}, z^{(n)}, r^{(n)}, h^{(n)}).$$

Абстрактна модель №4. Для четвертої абстрактної моделі суб'єкт першого речення частково виступає суб'єктом для всіх наступних. Тобто всі речення фрагменту розкривають одну тему, основи якої закладені у першому реченні тексту.

Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$L(S^{(1)}) = p^{(1)}(x_1^{(1)} \& \dots \& x_n^{(1)}, g^{(1)}, y^{(1)}, q^{(1)}, z^{(1)}, r^{(1)}, h^{(1)}) \&$$

$$L(S^{(2)}) = p^{(2)}(x_1^{(1)}, g^{(2)}, y^{(2)}, q^{(2)}, z^{(2)}, r^{(2)}, h^{(2)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(i)}) = p^{(i)}(x_i^{(1)}, g^{(i)}, y^{(i)}, q^{(i)}, z^{(i)}, r^{(i)}, h^{(i)}) \&$$

$$\dots \dots \dots$$

$$L(S^{(n)}) = p^{(n)}(x_n^{(1)}, g^{(n)}, y^{(n)}, q^{(n)}, z^{(n)}, r^{(n)}, h^{(n)}).$$

Абстрактна модель №5. У п'ятій моделі суб'єктами наступних речень можуть бути об'єкти, предмети та суб'єкти першого речення. У такому фрагменті передбачається відновлення відсутніх у тексті змістовних елементів шляхом їх заміни структурними компонентами логіко-лінгвістичних моделей першого речення тексту, що розглядається.

Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$L(S^{(1)}) = p^{(1)}(x^{(1)}, g^{(1)}, y^{(1)}, q^{(1)}, z^{(1)}, r^{(1)}, h^{(1)}) \&$$

$$L(S^{(2)}) = p^{(2)}(y^{(1)}, g^{(2)}, y^{(2)}, q^{(2)}, z^{(2)}, r^{(2)}, h^{(2)}) \&$$

$$L(S^{(i)}) = p^{(i)}(z^{(i)}, g^{(i)}, y^{(i)}, q^{(i)}, z^{(i)}, r^{(i)}, h^{(i)}) \&$$

$$L(S^{(n)}) = p^{(n)}(x^{(1)}, g^{(n)}, y^{(n)}, q^{(n)}, z^{(n)}, r^{(n)}, h^{(n)}).$$

Отже, послідовність розгортання змісту текстового фрагменту можна визначити за компонентами логіко-лінгвістичних моделей, зокрема, за зв'язками між їх суб'єктами, об'єктами та предметами відношень.

Обговорення результатів

Нехай розглядається фрагмент тексту: "...Відзначена закономірність побудови ланцюжків є регулярною, що дозволяє сформувати її аксіоматично у вигляді трьох постулатів. Перший постулат є цілком очевидним: кожне українське дієслово реалізується в мові лексемою. Другий постулат: кожна дієслівна лексема з визначеною і зафіксованою лексичною семантикою та конкретним значенням виду може реалізовуватися дієсловами. Третій постулат: кожна дієслівна лексема з конкретним значенням виду і зафіксованою лексичною семантикою може реалізовуватися дієсловами, які мають не більше чотирьох фонетичних варіантів..." [10]. Необхідно визначити його тематичну направленість.

Для кожного речення заданого текстового фрагменту будуватиметься логіко-лінгвістична модель, в результаті чого отримується множина логічних формул, які формально описують речення природної мови:

$L(S^{(1)}) = \epsilon$ & регулярною (закономірність, відзначена, побудови, 0, ланцюжків, 0) → дозволяє & сформувати (це, 0, закономірність, відзначену, вигляді & постулатів, трьох, аксіоматично).

$L(S^{(2)}) = \epsilon$ & очевидним (постулат, перший, 0, 0, 0, 0, цілком) → [реалізується (дієслово, кожне, лексемою, 0, 0, 0, мові)] & (реалізується (дієслово, українське, лексемою, 0, 0, 0, мові)].

$L(S^{(3)}) =$ Другий (постулат, 0, 0, 0, 0, 0, 0) → [(може & реалізовуватися (лексема, кожна, семантикою, визначеною, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, дієслівна, семантикою, визначеною, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, кожна, семантикою, зафіксованою, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, дієслівна, семантикою, зафіксованою, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, кожна, семантикою, лексичною, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, дієслівна, семантикою, лексичною, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, кожна, значенням & виду, конкретним, дієсловами, 0, 0)) &

(може & реалізовуватися (лексема, дієслівна, значенням & виду, конкретним, дієсловами, 0, 0))].

$L(S^{(4)}) =$ Третій (постулат, 0, 0, 0, 0, 0, 0) → [(може & реалізовуватися (лексема, кожна, значенням & виду, конкретним, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, дієслівна, значенням & виду, конкретним, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, кожна, семантикою, зафіксованою, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, дієслівна, семантикою, зафіксованою, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, кожна, семантикою, лексичною, дієсловами, 0, 0)) & (може & реалізовуватися (лексема, дієслівна, семантикою, лексичною, дієсловами, 0, 0))] → [(мають (дієслова, 0, варіантів, чотирьох, 0, 0, не & більше)) & (мають (дієслова, 0, варіантів, чотирьох, 0, 0, не & більше)) & (мають (дієслова, 0, варіантів, фонетичних, 0, 0, не & більше))].

Предмет відношень та його характеристика з першого речення фрагменту тексту "три постулати" виступають суб'єктом відношень та відношенням безпосереднього для наступних речень фрагменту: "перший постулат", "другий постулат", "третій постулат". Тому для даного фрагменту характерна четверта абстрактна модель, що відповідає тематичній прогресії з розщепленою ремою (новою інформацією по тексту). Відновлення компонентів логіко-лінгвістичних моделей, тобто відтворення відсутніх слів у реченнях природної мови, але присутніх у тексті змістовно, відбувається на основі використання принципів синтезу логіко-лінгвістичних моделей [11].

Висновки

У статті описано формальний апарат у вигляді логіко-лінгвістичних моделей представлення знань речень природної мови, що являється підґрунтям для виявлення логічних зв'язків між частинами текстів [12]. Текст представляється у вигляді множини логіко-лінгвістичних моделей, логічні операції та структурні компоненти яких вказують на типи логічних зв'язків.

Дослідження можливих способів побудови текстів довільної тематики дало змогу сформувати абстрактні моделі побудови логічних зв'язків у текстових фрагментах. Результатом побудови таких абстрактних моделей є багатогранники, що в подальшому використовуються для визначення тематики текстів.

Список літератури

- 1 Zhang, Y. Discriminative syntax-based word ordering for text generation / Y. Zhang // *Computational linguistics*. – 2015. – Vol. 41, №3. – P. 503-538. – doi:10.1162/COLI_a_00229.

- 2 **Evans, V.** Lexical concepts, cognitive models and meaning-construction / **V. Evans** // *Journal of Cognitive semiotics*. – 2006. – P. 73-107. – doi:10.1515/COG.2006.016.
- 3 **Gries, S. Th.** Corpus-based methods and cognitive semantics: the many meanings of to run / **S. Th. Gries** // *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*. – 2006. – P. 57-99.
- 4 **Jiawei, H.** Data Mining: Concepts and Techniques / **H. Jiawei, K. Micheline** // *Morgan Kaufmann Publishers*. – 2000. – 28 p.
- 5 **Jurafsky, D.** Natural Language Processing / **D. Jurafsky, C. Manning**. – Stanford University. – 2012. – [Web]: <https://www.coursera.org/course/nlp/> – 15.05.2014.
- 6 **Широков, В. А.** Корпусна лінгвістика / **В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна** та ін. – К.: Довіра, 2005. – 471 с.
- 7 **Лайонз, Дж.** Лингвистическая семантика. Монография / **Дж. Лайонз**. – М.: Языки славянской культуры. – 2003. – 400 с.
- 8 **Гальперин, И. Р.** Текст как объект лингвистического исследования. Изд. 5-тое, стереотипное / **И. Р. Гальперин**. – М: КомКнига. – 2007. – 144 с.
- 9 **Филиппов, К. А.** Лингвистика текста. Курс лекций / **К. А. Филиппов**. – СПб.: Издательство С.-Петербургского университета. – 2008. – 336 с.
- 10 **Широков, В. А.** Лінгвістичні та технологічні основи тлумачної лексикографії / **В. А. Широков, В. М. Білоноженко, О. В. Бугаков** та ін. – К.: Довіра. – 2010. – 295 с.
- 11 **Вавиленкова, А. И.** Основные принципы синтеза логико-лингвистических моделей / **А. И. Вавиленкова** // *Кибернетика и системный анализ*. – 2015. – Т. 51, № 5. – С. 176-185.
- 12 **Chang, C.** Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method / **C. Chang, S. Clark** // *Computational linguistics*. – 2014. – Vol. 40, №2. – P. 403-448. – doi:10.1162/COLI_a_00176.

Bibliography (transliterated)

- 1 **Zhang, Y.** Discriminative syntax-based word ordering for text generation. *Computational linguistics*, 2015, **41(3)**, 503-538, doi:10.1162/COLI_a_00229.
- 2 **Evans, V.** Lexical concepts, cognitive models and meaning-construction. *Journal of Cognitive semiotics*, 2006, 73-107, doi:10.1515/COG.2006.016.
- 3 **Gries, S. Th.** Corpus-based methods and cognitive semantics: the many meanings of to run. *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*, 2006, 57-99.
- 4 **Jiawei, H.** Data Mining: Concepts and Techniques, *Morgan Kaufmann Publishers*, 2000, 28 p.
- 5 **Jurafsky, D., Manning C.** Natural Language Processing, Stanford University, 2012. – [Web]: <https://www.coursera.org/course/nlp/> – 15.05.2014.
- 6 **Shirokov, V. A., Bugakov, O. V., Grjaznyxina, T. O.** Korpusna lingvistika, [Corpus linguistics], Kiev: *Dovira*, 471 p.
- 7 **Lyons, J.** Lingvisticheskaya semantika. Monografiya, [Linguistic semantics. Monograph], Moscow: *Yaziki slavyanskoj kylyri*, 2003, 400 p.
- 8 **Galperin, I. R.** Tekst kak ob"ekt lingvisticheskogo issledovaniya. [Text as an object of linguistic research], Moscow: *Komkniga Publ*, 2007, 144 p.
- 9 **Phillipov, K. A.** Lingvistika teksta, Sankt-Peterburg: *St. Petersburg university Publ.*, 2008, 336 p.
- 10 **Shirokov, V. A., Bilonozhenko, V. M., Bugakov, O. V.** Lingvistika teksta. Kurs lektсий. [Linguistic and technological bases of interpretative lexicography], Kiev, *Dovira Publ.*: 2010, 295 p.
- 11 **Vavilenkova, A. I.** Osnovnye printsipy sinteza logiko-lingvisticheskikh modeley [Basic Principles of the Synthesis of Logical-Linguistic Models], *Cybernetics and Systems Analysis*, 2015, **51(5)**, 826-834.
- 12 **Chang, C., Clark, S.** Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method, *Computational linguistics*, 2014, **40(2)**, 403-448, doi:10.1162/COLI_a_00176.

Відомості про авторів (About authors)

Вавиленкова Анастасія Ігорівна – кандидат технічних наук, доцент, Національний авіаційний університет, доцент кафедри комп'ютеризованих систем управління; м. Київ, Україна; e-mail: a_vavilenkova@mail.ua.

Anastasia Vavilenkova – Candidate of Technical Sciences (Ph. D.), Docent, Associate Professor, Department of Computerized Control Systems, National Aviation University, Kyiv, Ukraine; e-mail: a_vavilenkova@mail.ru.

Будь ласка посилайтесь на цю статтю наступним чином:

Вавиленкова, А. І. Способи виявлення логічних зв'язків між частинами текстових документів / **А. І. Вавиленкова** // *Вісник НТУ «ХПІ»*, Серія: Нові рішення в сучасних технологіях. – Харків: НТУ «ХПІ». – 2016. – № 12 (1184). – С. 101-105. – doi:10.20998/2413-4295.2016.12.14.

Please cite this article as:

Vavilenkova, A. Methods of identifying logical connections between parts of text documents. *Bulletin of NTU "KhPI"*. Series: *New solutions in modern technologies*. – Kharkiv: NTU "KhPI", 2016, **12** (1184), 101-105, doi:10.20998/2413-4295.2016.12.14.

Пожалуйста ссылайтесь на эту статью следующим образом:

Вавиленкова, А. И. Способы обнаружения логических связей между частями текстовых документов / **А. И. Вавиленкова** // *Вестник НТУ «ХПИ»*, Серія: Новые решения в современных технологиях. – Харьков: НТУ «ХПИ». – 2016. – № 12 (1184). – С. 101-105. – doi:10.20998/2413-4295.2016.12.14.

АННОТАЦІЯ Целью статьи является создание формального описания способов обнаружения логических связей между частями текстов. Одним из способов лингвистического анализа текстовой информации есть логико-лингвистические модели. Произвольный текст можно представить в виде массива логико-лингвистических моделей естественного языка, логические операции и структурные компоненты которых указывают на типы логических связей. Предлагается использовать для отображения структуры текста формальное описание пяти абстрактных моделей, что в будущем дает возможность определить тематику текста

Ключові слова: логико-лингвистические модели, естественный язык, логические связи, текстовый документ, анализ.

Надійшла (received) 21.03.2016