

КРИТЕРІЇ АНАЛІЗУ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ

А. І. ВАВІЛЕНКОВА

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

АНОТАЦІЯ Для здійснення змістовного аналізу електронних текстових документів запропоновано використовувати формальні логіко-лінгвістичні моделі. Метою статті є опис критеріїв аналізу формальних моделей, що здатні відображати зміст речень природної мови та формуються з використанням математичного апарату логіки предикатів. Описані критерії аналізу логіко-лінгвістичних моделей необхідні для побудови формальних моделей електронних текстових документів.

Ключові слова: логіко-лінгвістична модель; природна мова; аналіз; зміст; текст

ANALYSIS CRITERIA OF LOGIC AND LINGUISTIC MODELS OF NATURAL LANGUAGE SENTENCES

A. VAVILENKOVA

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

ABSTRACT The article describes the main text models used today as a tool for content processing electronic text documents. To make a content analysis author proposes to use formal logic and linguistic models, which are based on functional relationships between the principal and subordinate parts of natural language sentences. The article is to describe the criteria for analysis of formal models that can reflect the content of natural language sentences and which are formed using mathematical tools of predicate logic. For this purpose, the study researches principles of construction of logic and linguistic models of natural language sentences and formulates four criteria of analysis. First criterion analyzes the number of simple predicates in logic and linguistic model that helps to identify information about the type and composition of natural language sentences. The second criterion analyzes potency of set of predicate variables and constants of logic and linguistic model, which affects the number of simple predicates and identifies the type of individual forms of logic and linguistic model. The third criterion focuses on the analysis of logical operations that used in logic and linguistic model. That makes it possible to analyze the sequence of considerations referred to the natural language sentence. The forth criterion examines the presence of identical components in logic and linguistic models of natural language sentences from different sets of predicate variables and constants. Described analysis criteria of logic and linguistic models required to build formal models of electronic text documents using the mathematical apparatus of predicate logic.

Keywords: logic and linguistic model; natural language; analysis; content; text

Вступ

На сьогоднішній день існує безліч практичних задач, що потребують змістовного аналізу різноманітних текстових електронних документів [1]. Цілком очевидно, що вручну, тобто без застосування інформаційних технологій, вирішити ці задачі неможливо. Проте першочерговою проблемою, що стоїть на шляху вирішення задач змістовного аналізу є відсутність теорії побудови ефективних інструментів змістовної обробки електронних текстових документів. Одним із таких інструментів є моделі пошуку.

У наш час для вилучення і представлення певної інформації з електронних текстових документів у комп'ютерних технологіях використовуються декілька типів моделей [2]:

– векторно-просторова модель, коли документ представляється у вигляді вектора ваг, що обрані, як правило, за частотою вживання відповідних термів у документі [3];

– булева модель – базується на теорії множин та математичній логіці, вагові значення терму в документі приймають лише два значення: 0 або 1 [3];

– ймовірнісна модель передбачає наявність навчальних наборів релевантних та не релевантних документів, обраних користувачем або отриманих автоматично на певному тестовому реченні [4];

– семантична мережа – розглядається як множина понять, в яку включаються слова тексту, що найбільш часто вживаються, а відповідно несуть найбільше змістовне навантаження; для кожного поняття формується набір асоціативних зв'язків, і чим частіше зустрічаються два поняття в реченнях тексту, тим вища ймовірність того, що вони пов'язані за змістом [5];

– мережа природних ієрархій термінів – алгоритм її формування передбачає реалізацію послідовності кроків, що охоплюють попередню обробку початкового текстового корпусу, визначення та сортування термінів, вибір необхідної кількості

найбільш вагомим вузлів графу горизонтального бачення [6].

При побудові більшості моделей пошуку зміст визначається множиною слів та стійких словосполучень – термів, які входять до нього без врахування взаємозв'язків, як "мішок зі словами" [7]. Таке припущення приводить до втрати змістовних відтінків, проте дозволяє реалізувати пошук та здійснити групування документів за формальними ознаками.

Тому важливим моментом для змістовного аналізу електронних текстових документів є правильний вибір формальної моделі представлення знань, що здатна була б відображати зміст текстової інформації.

Мета роботи

Метою роботи є опис критеріїв аналізу формальних моделей, що здатні відображати зміст речень природної мови та формуються з використанням математичного апарату логіки предикатів. Досягнення поставленої мети передбачає:

- аналіз існуючих текстових моделей;
- дослідження принципів побудови логіко-лінгвістичних моделей речень природної мови;
- формулювання критеріїв аналізу логіко-лінгвістичних моделей речень природної мови;
- демонстрація результатів застосування описаних критеріїв аналізу до речень природної мови з текстових документів наукового стилю.

Викладення основного матеріалу

Речення – це мінімальна і основна комунікативна одиниця мови, що має бути цілісною і передавати інформацію в усій складності залежностей і зв'язків [8]. Тому в основу принципів побудови логіко-лінгвістичних моделей покладено функціональні залежності між головними та другорядними членами речень природної мови. Керуючись гіпотезою "правило на правило", згідно з якою між граматичною структурою речення та логічною формою існує повна відповідність, можливістю застосування теореми про повноту [9], а також враховуючи те, що кожне речення природної мови характеризується граматичною організацією: будовою та граматичним вираженням (оформленням) членів речення, маємо:

- підмет – суб'єкт;
- присудок – відношення;
- додаток – об'єкт та предмет відношення;
- означення – характеристики суб'єкта та об'єкта та предмет відношення;
- обставина – характеристика відношення.

Таким чином, побудова логіко-лінгвістичних моделей речень природної мови [10] відбувається за наступними принципами:

1) визначення кількості простих атомарних предикатів v^s ;

2) фіксація значень елементів кортежу логічних операцій $O(S)$ речення S :

$$O(S) = [o_e(S) \mid e = \overline{1, h}],$$

де h – загальна кількість логічних операцій, наявних у реченні S ;

3) визначення множини відношень P^s між суб'єктами та об'єктами речення S , що фактично містить дії, стани або властивості суб'єкта;

4) фіксація множини предикатних констант H_p^s , що містять характеристики відношень між суб'єктами та об'єктами у реченні S ;

5) визначення множини суб'єктів $X_p^s(h)$, що входять до речення S та пов'язані з об'єктами речення p -им відношенням, яке володіє h -ю характеристикою;

6) фіксація множини предикатних констант $G_p^s(x, h)$, що містять характеристики (параметри) суб'єктів, які входять до речення S ;

7) визначення множини об'єктів $Y_p^s(x, g, h)$, що входять до речення S та пов'язані з суб'єктами речення $x \in X_p^s(h)$ з характеристиками $g \in G_p^s(x, h)$ p -им відношенням, яке володіє h -ю характеристикою;

8) фіксація множини предикатних констант $Q_p^s(x, g, y, h)$, що містять характеристики (параметри) об'єктів, які входять до речення S ;

9) визначення множини предметів $Z_p^s(x, g, y, q, h)$ p -го відношення, яке володіє h -ю характеристикою, між суб'єктами $x \in X_p^s(h)$ з характеристиками $g \in G_p^s(x, h)$ та об'єктами $y \in Y_p^s(x, g, h)$ з характеристиками $q \in Q_p^s(x, g, y, h)$;

10) фіксація множини предикатних констант $R_p^s(x, g, y, q, z, h)$, що містить характеристики (параметри) предметів p -го відношення, яке володіє h -ю характеристикою, між суб'єктами $x \in X_p^s(h)$ з характеристиками $g \in G_p^s(x, h)$ та об'єктами $y \in Y_p^s(x, g, h)$ з характеристиками $q \in Q_p^s(x, g, y, h)$ у реченні S ;

11) формування логічного виразу L^s , шляхом підстановки у нього змінних, отриманих за попередніми принципами побудови логіко-лінгвістичних моделей відповідно до кількості елементів зафіксованих множин.

Логіко-лінгвістичні моделі, побудовані за описаними вище принципами, є інтерпретацією семантико-синтаксичної структури речень природної

мови, так як будуються на основі наявних у реченні означальних (між означенням і підметом, означенням і додатком), обставинних (між присудком і обставинами), додаткових (між присудком і додатком), присудкових (між підметом і присудком) зв'язків.

Керуючись принципами побудови логіко-лінгвістичних моделей речень природної мови, можна сформулювати критерії для їх аналізу.

Критерій 1. Кількість простих предикатів в логіко-лінгвістичній моделі речення природної мови L^S . Цей критерій дає змогу визначити інформацію про тип та склад речення природної мови.

Якщо кількість простих предикатів в логіко-лінгвістичній моделі рівна одному $v^S = 1$, тобто $L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1)$, то вона описує просте речення природної мови без однорідних членів.

Якщо кількість простих предикатів більше одного $v^S > 1$ і при цьому прості предикати відрізняються однією або декількома (одна з яких відповідає за характеристику) компонентами, то логіко-лінгвістична модель описує просте речення з однорідними членами. Наприклад, логіко-лінгвістична модель

$L^S = p_{11}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1)$ описує речення з однорідними присудками, а модель $L^S = p_1(x_{11}, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_1(x_{11}, g_1, y_1, q_{12}, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_{12}, z_1, r_1, h_1)$ описує просте поширене речення з однорідними підметами та однорідними означеннями, що характеризують об'єкт відношення.

Якщо кількість простих предикатів більше одного $v^S > 1$ і при цьому майже всі предикатні змінні та константи різні, то логіко-лінгвістична модель описує складне речення, яке в подальшому необхідно проаналізувати за ще декількома критеріями. Наприклад,

$$L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p'_1(x'_1, g'_1, y'_1, q'_1, z'_1, r'_1, h'_1).$$

Критерій 2. Потужність множин предикатних змінних та констант логіко-лінгвістичної моделі.

Потужність множин предикатних змінних та констант впливає, в першу чергу, на кількість простих предикатів v^S :

$$v^S = \sum_{p \in P^S} \sum_{h \in H_p^S} v_p^S(h),$$

$$v_p^S(h) = \sum_{x \in X_p^S(h)} \sum_{g \in G_p^S(x, h)} v_p^S(x, g, h),$$

$$v_p^S(x, g, h) = \sum_{y \in Y_p^S(x, g, h)} \sum_{q \in Q_p^S(x, g, y, h)} v_p^S(x, g, y, q, h),$$

$$v_p^S(x, g, y, q, h) = \sum_{z \in Z_p^S(x, g, y, q, h)} |R_p^S(x, g, y, q, z, h)|.$$

Також потужність множин визначає тип окремої форми логіко-лінгвістичної моделі.

Якщо просте речення природної мови містить лише одну частину, що має закінчений зміст, тобто $|P^S|=1$; $|H_p^S|=1$; $|X_p^S(h)|=1$; $|G_p^S(x, h)|=1$; $|Y_p^S(x, g, h)|=1$; $|Q_p^S(x, g, y, h)|=1$; $|Z_p^S(x, g, y, q, h)|=1$; $|R_p^S(x, g, y, q, z, h)|=1$.

Якщо множина H_p^S характеристик хоча б одного p -го відношення у реченні S між суб'єктом x та об'єктом y містить більше одного елемента і $|X_p^S(h)|=1$; $|G_p^S(x, h)|=1$; $|Y_p^S(x, g, h)|=1$; $|Q_p^S(x, g, y, h)|=1$; $|Z_p^S(x, g, y, q, h)|=1$; $|R_p^S(x, g, y, q, z, h)|=1$, то у реченні природної мови йдеться про різні обставини дії.

Якщо множина суб'єктів $X_p^S(h)$, що пов'язані з об'єктами речення p -им відношенням, яке володіє h -ю характеристикою, містить декілька елементів і $|P^S|=1$; $|H_p^S|=1$; $|G_p^S(x, h)|=1$; $|Y_p^S(x, g, h)|=1$; $|Q_p^S(x, g, y, h)|=1$; $|Z_p^S(x, g, y, q, h)|=1$; $|R_p^S(x, g, y, q, z, h)|=1$, то речення просте з однорідними підметами.

Якщо множина об'єктів $Y_p^S(x, g, h)$, що пов'язані з суб'єктами речення p -им відношенням, яке володіє h -ю характеристикою, містить декілька елементів і $|P^S|=1$; $|H_p^S|=1$; $|X_p^S(h)|=1$; $|G_p^S(x, h)|=1$; $|Q_p^S(x, g, y, h)|=1$; $|Z_p^S(x, g, y, q, h)|=1$; $|R_p^S(x, g, y, q, z, h)|=1$, то у реченні йдеться про декілька об'єктів відношень.

Якщо потужність множин $|P^S| \geq 2$; $|X_p^S(h)| \geq 2$, то речення природної мови складне, а v^S завжди більше одиниці.

Критерій 3. Логічні операції, що використані в логіко-лінгвістичній моделі речення природної мови L^S . Так:

– операція кон'юнкції «&» свідчить про наявність у реченні послідовності певних дій, перерахування, ствердження істинності простих предикатів, що з'єднані цією операцією;

– операція диз'юнкції « \vee » свідчить про наявність у реченні послідовності подій, що протиставляються, зіставляються, взаємно виключаються або чергуються;

– операція імплікації « \rightarrow » використовується між реченнями природної мови, у яких йдеться про певну подію та наслідок (результат) її виконання.

Таким чином, логічні операції дають змогу проаналізувати послідовність розмірковувань, про які йдеться у реченні природної мови.

Критерій 4. Наявність тотожних компонентів у логіко-лінгвістичній моделі речення природної мови із різних множин предикатних змінних та констант.

Якщо тотожні елементи знайдені в різних множинах однієї і тієї ж логіко-лінгвістичної моделі, то це свідчить про наявність у реченні синонімів, займенників або вказівних слів, що дозволили зробити в моделі синтез [11]. Аналіз того, в яких саме множинах предикатних змінних або констант наявні такі елементи дасть можливість говорити про те, які прийоми змістовної заміни були зроблені у реченні природної мови.

Наприклад, якщо збігаються елементи множини суб'єктів та предметів відношень у двох різних простих предикатах, то розглядається складне означальне речення, у якому останнє слово першого речення виступає суб'єктом другого, залежного речення, вираженим займенником.

Обговорення результатів

Сформульовані критерії аналізу електронних текстових документів пропонується застосувати для визначення та відновлення максимальної інформації з тексту наукового стилю [12], представленого такою логіко-лінгвістичною моделлю:

$$L^S = p_1(0,0, y_1, q_1, z_1, r_1, h_1) \& (p'_1(h_1,0, y'_1,0, z'_1, r'_{11},0) \& p'_1(h_1,0, y'_1,0, z'_1, r'_{12},0) \& p'_1(h_1,0, y'_1,0, z'_1, r'_{13},0)).$$

$L^S =$ *Описано(0,0,процес,формування,бази, знань&СППР, обставин) & (пов'язані(обставини,0,виявленням,0,ознак, важко&пояснювальних,0) & (пов'язані(обставини,0,виявленням,0,ознак, аномалій,0) & (пов'язані(обставини,0,виявленням,0,ознак,атак,0)).*

Логіко-лінгвістична модель послідовно піддається аналізу за кожним із сформульованих критеріїв.

За першим критерієм визначається кількість простих предикатів в логіко-лінгвістичній моделі речення природної мови, у даному випадку $v^S = 4$, що говорить про те, що речення не являється простим. Три прості предикати, що входять до складу моделі містять всі однакові компоненти, крім характеристики предмета відношень, четвертий простий предикат містить зовсім інші значення компонент. Це означає, що речення природної мови складне, а його друга частина містить однорідні означення.

За другим критерієм аналізу виявлено, що потужності множин відношень та суб'єктів відношень містять дорівнюють двом, а потужність множини характеристик предмета відношень дорівнює чотирьом. Отже, ще раз підтверджено, що речення природної мови складне.

За третім критерієм аналізу визначено, що у реченні йдеться про послідовність певних дій, перерахування, ствердження істинності простих предикатів, так як в логіко-лінгвістичній моделі використано лише операцію кон'юнкції.

Після застосування четвертого критерію аналізу видно, що в логіко-лінгвістичній моделі тотожними є елементи з множини характеристик відношень та множини суб'єктів, що говорить про використання у реченні дієприслівникового звороту або підрядного означального зв'язку.

Якщо співставити результати, отримані в результаті застосування критеріїв аналізу, та оригінальний текст: «*Описано процес формування бази знань СППР для обставин, пов'язаних з виявленням важко пояснюваних ознак аномалій та атак*» [12], то видно, що критерії аналізу дали достатньо точні показники щодо структури речення природної мови.

Висновки

У статті описано основні принципи формування логіко-лінгвістичних моделей, побудованих на основі використання граматичної організації речень природної мови. Це дало можливість сформулювати чотири критерії аналізу таких логіко-лінгвістичних моделей, за допомогою яких можна визначити інформацію про тип та склад речення природної мови та проаналізувати послідовність розмірковувань, про які йдеться у реченні.

Логіко-лінгвістична модель текстового документу – це абстрактна модель, яка об'єднує в собі основні властивості тексту та його складових частин, відображає основні взаємозв'язки між структурними компонентами, представляє собою впорядковану четвірку та масив логіко-лінгвістичних моделей речень природної мови, що входять до тексту. Тому описані критерії аналізу логіко-лінгвістичних моделей необхідні для побудови формальних моделей електронних текстових документів і являються необхідними складовими процесу їх автоматичного лінгвістичного аналізу.

Список літератури

- 1 **Jurafsky, D.** Natural Language Processing / **D. Jurafsky, C. Manning.** – Stanford University, 2012. – Available at: <https://www.coursera.org/course/nlp/> – 15.05.2014.
- 2 **Ландэ, Д. В.** Интернетика: навигация в сложных сетях: модели и алгоритмы / **Ландэ Д. В., Снарский А. А., Безсуднов И. В.** – М.: Либроком, 2009. – 264 с.
- 3 **Zhang, Y.** Discriminative syntax-based word ordering for text generation / **Y. Zhang** // *Computational linguistics.* – 2015. – Vol.41, №3. – P. 503–538.
- 4 **Evans, V.** Lexical concepts, cognitive models and meaning-construction / **V. Evans** // *Journal of Cognitive semiotics.* – 2006. – P. 73–107.
- 5 **Kumar, J.** Duplicate and near duplicate documents detection / **J. Kumar, P. Govindarajulu** // *European journal of scientific research.* – 2009. – Vol.32. – P. 514 – 527.
- 6 **Geeraerts, D.** Cognitive linguistics: basic readings research / **Geeraerts D., Dirven R., John R.** – Berlin-New York: Publ. mouton de cruyter, 2006. – 486 p.

- 7 **Jiawei, H.** Data Mining: Concepts and Techniques / **H. Jiawei, K. Micheline** // *Morgan Kaufmann Publishers*. – 2000.
- 8 **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.
- 9 **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.
- 10 **Вавіленкова, А. І.** Способи виявлення логічних зв'язків між частинами текстових документів / **А. І. Вавіленкова** // *Вісник Національного технічного університету «Харківський політехнічний університет»*: зб. наук. праць. – (Серія «Нові рішення в сучасних технологіях»). – 2016. – № 12 (1184). – С. 101 – 105. doi: 10.20998/2413-4295.2016.12.14.
- 11 **Вавіленкова, А. И.** Основные принципы синтеза логико-лингвистических моделей / **А. И. Вавіленкова** // *Кибернетика и системный анализ*. – 2015. – Т. 51, № 5. – С. 176 – 185.
- 12 **Akhmetov, B.** Support system for the weakly formalized problems in the provision of cybersecurity / **B. Akhmetov, V. Lakhno, Y. Boiko, A. Mishchenko** // *Eastern-European Journal of Enterprise Technologies*. – 2017. – № 1/2 (85). – P. 4 – 15, doi: 10.15587/1729-4061.2017.90506.
- 3 **Zhang, Y.** Discriminative syntax-based word ordering for text generation. *Computational linguistics*, 2015, Vol. **41**, №3, P. 503–538.
- 4 **Evans, V.** Lexical concepts, cognitive models and meaning-construction. *Journal of Cognitive semiotics*, 2006, P. 73–107.
- 5 **Kumar, J., Govindarajulu, P.** Duplicate and near duplicate documents detection. *European journal of scientific research*, 2009, Vol. **32**, P. 514 – 527.
- 6 **Geeraerts, D., Dirven, R., John, R.** Cognitive linguistics: basic readings research, Berlin-New York: Publ. mouton de cruyter, 2006, 486 p.
- 7 **Jiawei, H., Micheline, K.** Data Mining: Concepts and Techniques. *Morgan Kaufmann Publishers*, 2000.
- 8 **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, P. 57–99.
- 9 **Chang, C., Clark, S.** Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method. *Computational linguistics*, 2014, Vol. **40**, №2, P. 403–448.
- 10 **Vavilenkova, A. I.** 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.
- 11 **Vavilenkova, A. I.** Basic principles of synthesis of logic and linguistic models. *Cybernetics and Systems Analysis*, 2015, **51**, № 5, 176 – 185.
- 12 **Akhmetov, B., Lakhno, V., Boiko, Y., Mishchenko, A.** Support system for the weakly formalized problems in the provision of cybersecurity. *Eastern-European Journal of Enterprise Technologies*, 2017, № **1/2** (85), P. 4 – 15, doi: 10.15587/1729-4061.2017.90506.

Bibliography (transliterated)

- 1 **Jurafsky, D., Manning, C.** Natural Language Processing, Stanford University, 2012, Available at: <https://www.coursera.org/course/nlp/> – 15.05.2014.
- 2 **Lande, D. V., Snarskiy, A. A., Bezsudnov, I. V.** Internet: navigation in complex networks: models and algorithms, M.: Librocom, 2009, 264 c.

Відомості про авторів (About authors)

Вавіленкова 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:

Вавіленкова, А. І. Критерії аналізу логіко-лінгвістичних моделей речень природної мови / **А. І. Вавіленкова** // *Вісник НТУ «ХПІ»*, Серія: *Нові рішення в сучасних технологіях*. – Харків: НТУ «ХПІ». – 2017. – № 7 (1229). – С. 118-122. – doi:10.20998/2413-4295.2017.07.16.

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", 2017, **7** (1229), 118–122, doi:10.20998/2413-4295.2017.07.16.

Пожалуйста, ссылайтесь на эту статью следующим образом:

Вавіленкова, А. И. Критерии анализа логико-лингвистических моделей предложений естественного языка / **А. И. Вавіленкова** // *Вестник НТУ «ХПИ»*, Серія: *Новые решения в современных технологиях*. – Харьков: НТУ «ХПИ». – 2017. – № 7 (1229). – С. 118-122. – doi:10.20998/2413-4295.2017.07.16.

АННОТАЦИЯ Для осуществления содержательного анализа электронных текстовых документов предложено использовать формальные логико-лингвистические модели. Целью статьи является описание критериев анализа формальных моделей, которые способны отображать содержание предложений естественного языка, используя математический аппарат логики предикатов. Описаны критерии анализа логико-лингвистических моделей, необходимые для построения формальных моделей электронных текстовых документов.

Ключевые слова: логико-лингвистическая модель; естественный язык; анализ; содержание; текст

Поступила (received) 17.03.2017