Regularity of Context Units Identification in Electronic Text Documents

Anastasiia Vavilenkova

National Aviation University, Liubomyra Huzara ave,1, Kyiv, 03058, Ukraine

Abstract

In this article had been analyzed actual software services, that can build relation's tree and make syntactical analysis. Each of them transforms primary text into the data structure with special features. But none of them can identify context units from the sentence of natural language. The author proposed to use components of logic and linguistic models for automatic generation of grammar colocations. Also author suggested the rules for context units identification for complex sentences of natural language. It had been demonstrated the outcome of using these rules. It is a program, that extract all collocations from different types of natural language sentences.

Keywords [1]

Context units, identification, analysis, knowledge, syntactical parser, logic and linguistic model.

1. Introduction

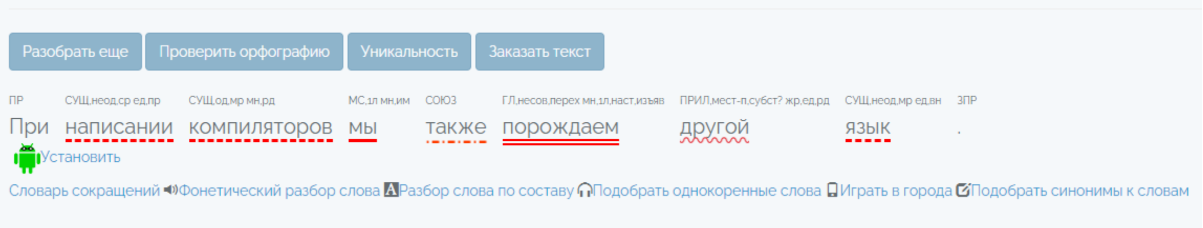
All early efforts to extract knowledge from the textual information by difference scientists leads in grammars by Homskiy and transformation grammars [1, 12]. Grammars of regularity do not come up with collocations like analysis ones. However, almost all linguistic theories describe linear sequence of the sentence units by mean of hierarchic structure of gramma components [2].

R. Shenk proposed an approach, that based on using equal concept constructions for identity sentences [3]. It assumes emergency of elementary situations and making a great number of templates.

Inevitably, it was not effectively used method for knowledge representation. Ch. Filmor offered the system of semantic roles, reflected logical structure, not only grammar. Semantic relations in this case are between context of the verb and context of noun group [13].

Nowadays many creators of analytical systems use APIs, that make the process of syntactical analysis easier. After syntactical parsing primary text transformed into the data structure for future processing.It had been analyzed some of these syntactical parsers.

The results of on-line transformer progaonline.com [18] work reflects on Figure 1.



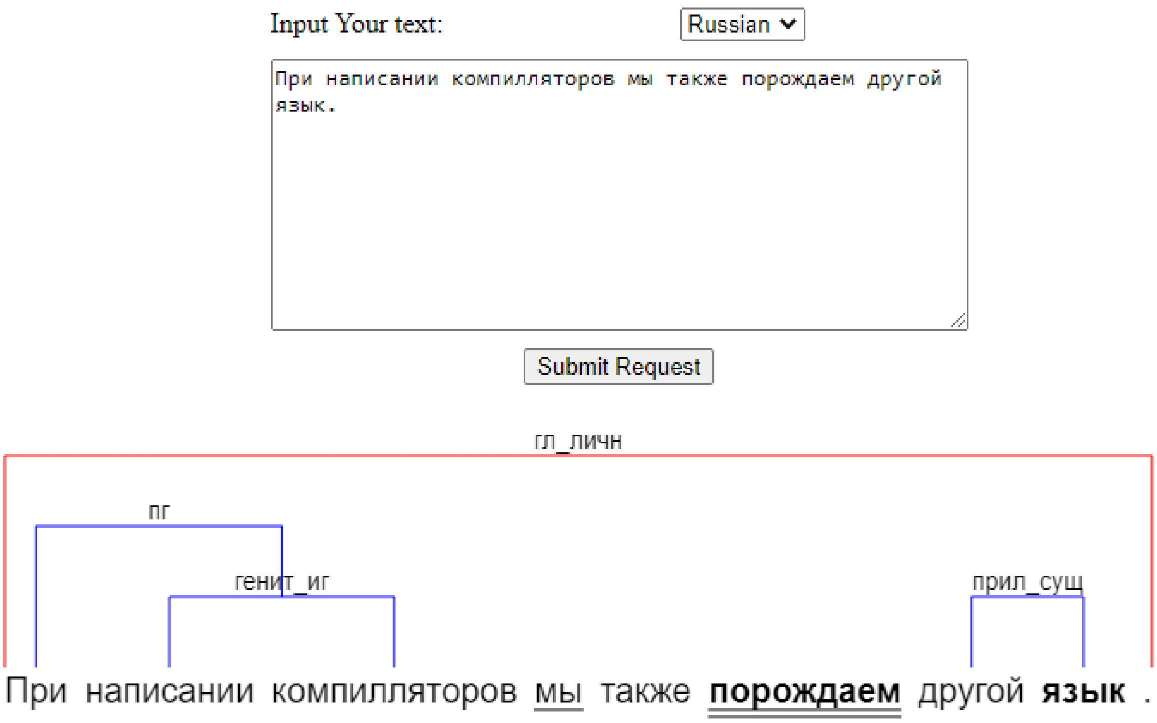
```
ПР   СУЩ,неод,ср ед,пр  СУЩ,од,мр мн,рд  МС,1л мн,им  СОЮЗ   ГЛ,несов,перех мн,1л,наст,изъяв  ПРИЛ,мест-п,субст? жр,ед,рд  СУЩ,неод,мр ед,вн  ЗПР
При  написании          компиляторов     мы           также  порождаем                        другой                       язык               .
```

Установить
Словарь сокращений
Фонетический разбор слова
Разбор слова по составу
Подобрать однокоренные слова
Играть в города
Подобрать синонимы к словам

Figure 1: The results of progaonline.com transformer work

System Aot.tu [19] transforms Russian and Germany sentences into relation's tree (Figure 2).

[1] *IT&I-2020 Information Technology and Interactions, December 02–03, 2020, KNU Taras Shevchenko, Kyiv, Ukraine*
EMAIL: vavilenkovaa@gmail.com (A. Vavilenkova)
ORCID: 0000-0002-9630-4951 (A. Vavilenkova)



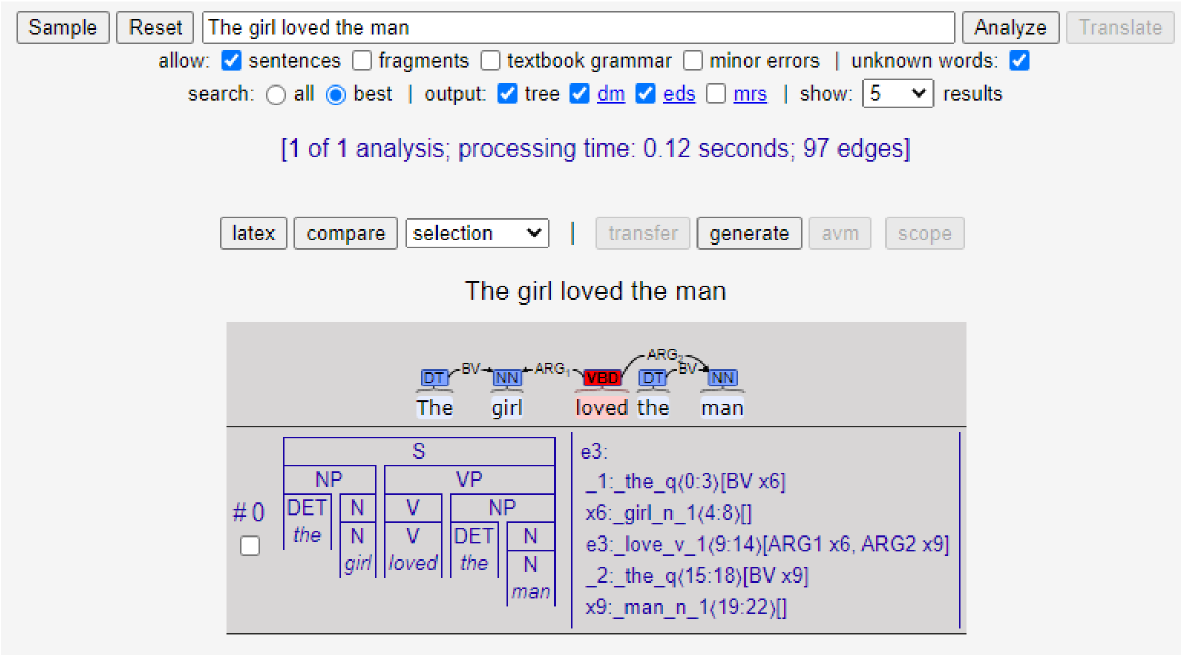
Input Your text: Russian

При написании компилляторов мы также порождаем другой язык.

Submit Request
гл_личн
пг
генит_иг
прил_сущ
При написании компилляторов мы также порождаем другой язык .

Figure 2: The results of Aot.ru system's work

Demo version of abroad open online parser erg.delph-in [20] builds a relation's tree for simple sentence according to the Filmor's scheme: each world in the sentence relate to the verb "love" (Figure 3).



The girl loved the man

allow: ☑ sentences □ fragments □ textbook grammar □ minor errors | unknown words: ☑

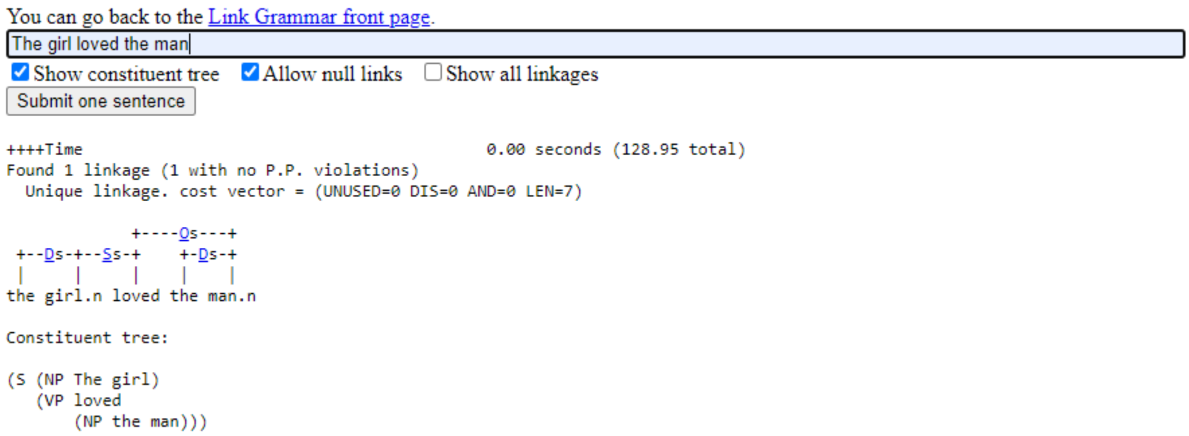
search: ○ all ● best | output: ☑ tree ☑ dm ☑ eds □ mrs | show: 5 results

[1 of 1 analysis; processing time: 0.12 seconds; 97 edges]

```
e3:
_1:_the_q⟨0:3⟩[BV x6]
x6:_girl_n_1⟨4:8⟩[]
e3:_love_v_1⟨9:14⟩[ARG1 x6, ARG2 x9]
_2:_the_q⟨15:18⟩[BV x9]
x9:_man_n_1⟨19:22⟩[]
```

Figure 3: The results of syntactical parser erg.delph-in work

The results of syntactical analysis by Link Grammar [21] for the same sentence are on Figure 4. As we can see, there are no open parsers for Ukrainian language. Parsers appropriate for transforming only Russian or English language. The results of transforming the same sentences are different for all testing parsers. Not surprisingly, traditional ways of analysis by key parameters and standard answers are not possible to analyze natural language text at all and are not helpful for context analysis [4–5].



You can go back to the Link Grammar front page.

The girl loved the man

☑ Show constituent tree ☑ Allow null links □ Show all linkages

Submit one sentence

```
++++Time                                          0.00 seconds (128.95 total)
Found 1 linkage (1 with no P.P. violations)
  Unique linkage. cost vector = (UNUSED=0 DIS=0 AND=0 LEN=7)

             +----Os---+
 +--Ds-+--Ss-+    +-Ds-+
 |     |     |    |    |
the girl.n loved the man.n

Constituent tree:

(S (NP The girl)
   (VP loved
       (NP the man)))
```

Figure 4: The results of Link Grammar work

Today the main aim of natural language researches is automatic creation of context data structures for formalization of logical links by mean of particular algebraic construction. From the other hand, these researches give practical value for automatic analysis and synthesis of natural language texts by computer technologies.

The article is the result of the author's research in the text linguistics and formal semantics, combined with mathematical apparatus of first order predicate logic.

2. Materials and Methods

The most popular automatic operations for concept processing of text are searching, word's definition, abstracting and translation. All these types suppose searching and detection some regular expressions or textual fragments, that fulfilled some conditions. For instance, such type of detection is possible for recognizing telephone numbers, addresses and template's elements. Is this case, definition is realized by means of symbol comparison. The second treatment of textual fragments searching and detection is based on previously created database of indexes. For example, there are many bases with informal fragments, special thesauruses [14]. Despite almost a century of research in artificial intelligence, context units identification still can not be realized in correct form for complex sentences.

One of the most essential thing for all systems that woks with natural languages must be using the process of grammatical analysis [15]. Grammar is a set of rules expressed relations between the members of a sentence.

F. George [12] proposed the scheme on Figure 5 for interpretation the simple sentence "Jack hit the ball".

The variables *N, V, NP, VP, A, Adj, AP, AdjP* on the scheme are used for designation of noun, verb, group of subject, group of predicate, adjective, participle, group of participle and group of adjective and ect.

For more complex sentence "The large black dog fiercely chased the small boy away from the house" F. Jorge proposed the scheme on the Figure 6 and the syntactical graph on the Figure 7.

We can see different representation of the same content and any rules for that. As the author said, the number of variables can be increase. For instance, composite sentence "The fine silver veil fluttered from the shoulders of the dancer as if a summer breeze were blowing the shadow of clouds away from the white town, and came to rest on the dark ground" [12] will be represented as syntactical graph on Figure 8.

So the question "how to formalize textual fragments with a set of complex, logically connected sentences", come up and must be responded. Almost all methods of automatic analysis of natural language texts divide sentences on group of words according to the words' places, not according to the context. As distinct from these procedures of natural language sentence dividing, logic and linguistic modelling based on idea of collocations identification or finding context units. There is a full correspondence between grammar structure and logic form of natural language sentence of natural language sentence [6–7].

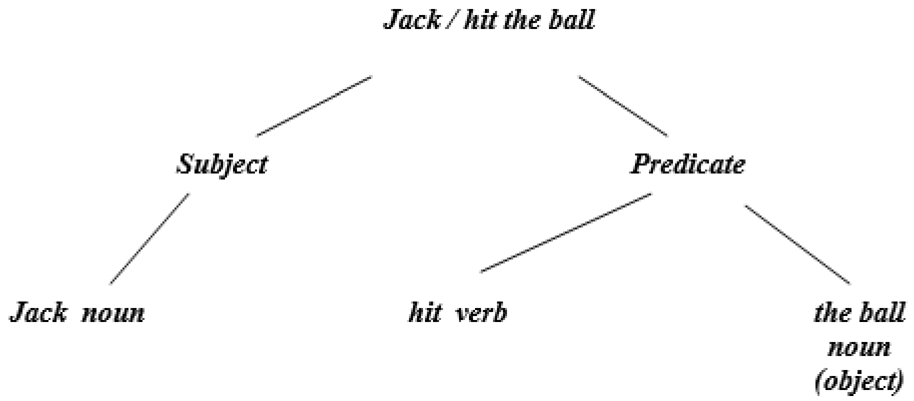


Figure 5: Interpretation of the simple sentence

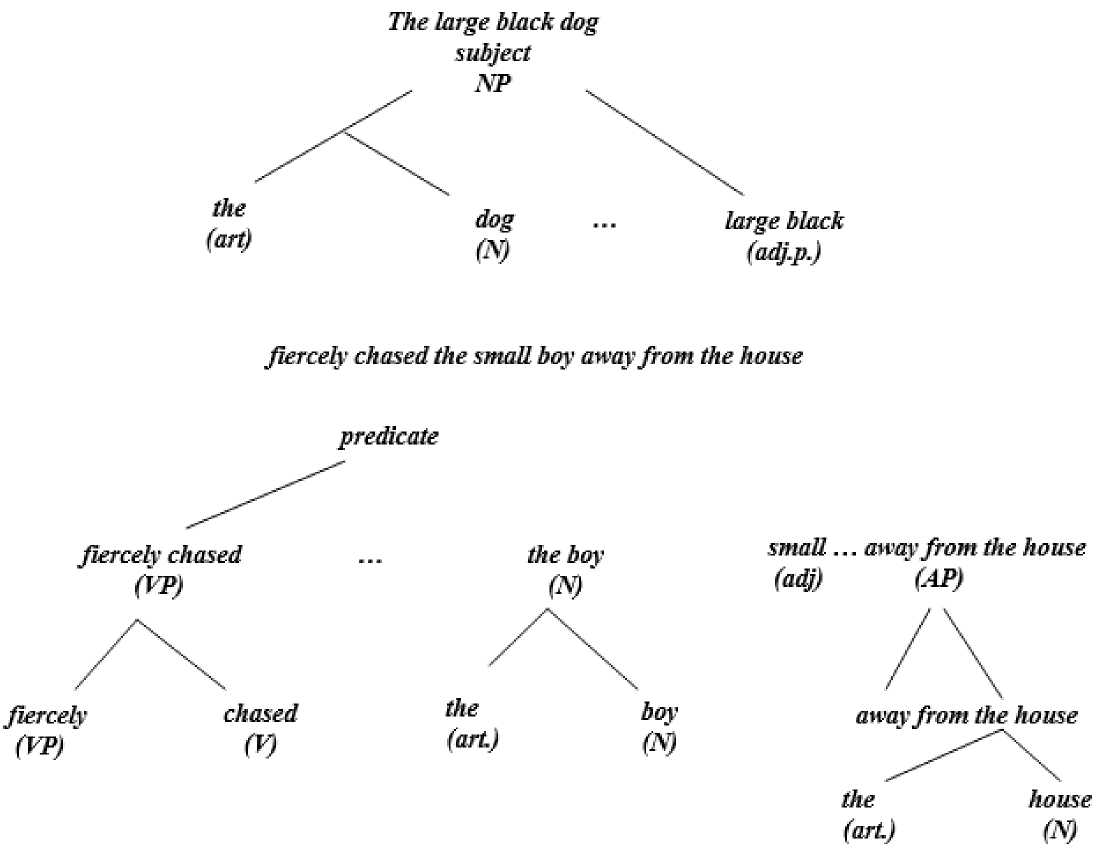


Figure 6: Interpretation of the complex sentence by F. Jorge

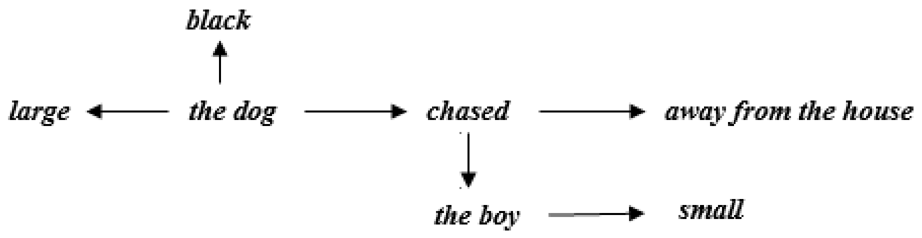


Figure 7: Example of syntactical graph

Considering grammar organization of the sentences, we have such graduation of sentences' members [8]:

- subject of the sentence – subject x;
- predicate of the sentence – relation p;
- object of the sentence – object y or subject-matter of relation z;
- definition – characteristic of subject g, characteristic of object q or characteristic of subject-matter of relation r;
- circumstance – characteristic of relation h.

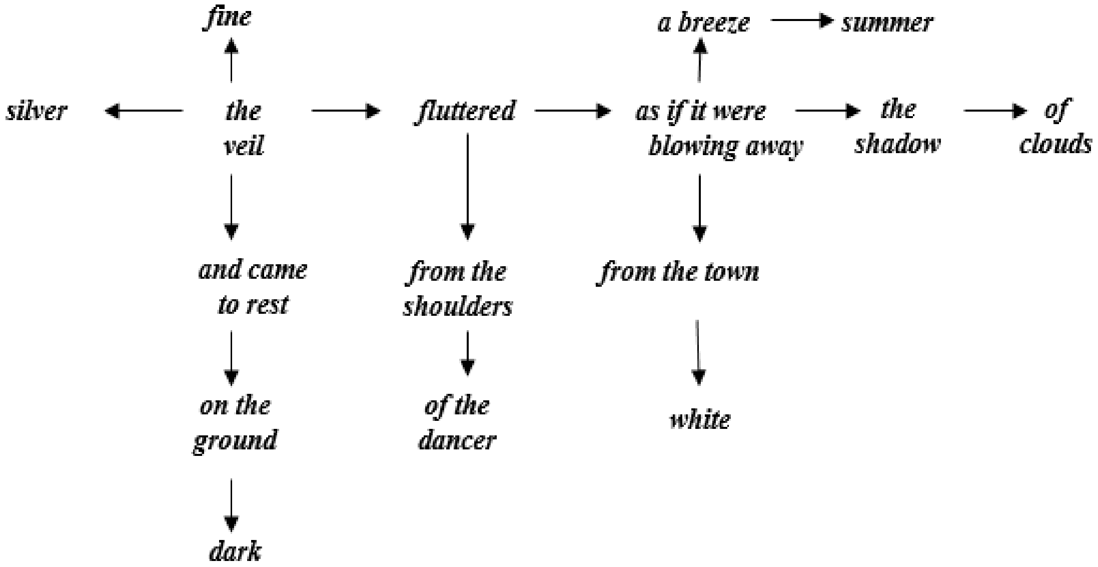


Figure 8: Example of syntactical parsing of composite sentence

A set of words connected between each other by logic links, will be lettered $sp_j, j=\overline{1,m}$, where m – amount of the collocations in the sentence.

According to the Ukrainian and English language rules, collocations can be formed between those members of sentence [9, 16]:

- "definition – subject" – $sp_j = g \cup x$;
- "predicate – object" – $sp_j = p \cup y$;
- "definition – object" – $sp_j = q \cup y$;
- "object – object" – $sp_j = y \cup z$;
- "object – object" – $sp_j = r \cup z$;
- "circumstance – predicate" – $sp_j = h \cup p$.

For instance, we need to detect collocations in the simple sentence "*Математичне моделювання часто використовують у навчальному процесі*". Firstly, roles for each world in the sentence are identified:

- subject of the sentence – subject x – *0*;
- predicate of the sentence – relation p – *використовують*;
- object of the sentence – object y – *моделювання*;
- object of the sentence – subject-matter of relation z – *процесі*;
- definition – characteristic of subject g – *0*;
- definition – characteristic of object q – *математичне*;
- definition – characteristic of subject-matter of relation r – *навчальних*;
- circumstance – characteristic of relation h – *часто*.

Thus, there are next collocations in this sentence [17]:

– “predicate – object” – $sp_j = p \cup y$ – *використовують моделювання*;

– “definition – object” – $sp_j = q \cup y$ – *математичне моделювання*;

– “object – object” – $sp_j = y \cup z$ – *процесі моделювання*;

– “definition – object” – $sp_j = r \cup z$ – *навчальному процесі;*

– “circumstance – predicate” – $sp_j = h \cup p$ – *часто використовують*.

Collocations for the similar English sentence “*Mathematical modelling often used in educational process*”:

– “predicate – object” – $sp_j = p \cup y$ – *used modelling*;

– “definition – object” – $sp_j = q \cup y$ – *mathematical modelling*;

– “object – object” – $sp_j = y \cup z$ – *modelling process*;

– “definition – object” – $sp_j = r \cup z$ – *educational process;*

– “circumstance – predicate” – $sp_j = h \cup p$ – *often used.*

Every world in the sentence S_i ($i = \overline{1,n}$, n – the amount of words in the sentence) can be represented by the set of characteristics:

$$Z_i(S_i) = \{cm_i, g_i, n_i, k2_i, t_i, h_i, l_i, ch_i\},$$

where $cm_i = \overline{1,11}$ – grammar characteristic, that means a part of speech, each number is responsible for one part of speech: noun – 1, adjective – 2, numeral – 3, pronoun – 4, verb –5, participle I – 6, participle II – 7, adverb – 8, preposition – 9, union –10 or particle – 11;

$g_i = \overline{1,7}$ – morphologic characteristic, that is responsible for case;

$n_i = \overline{1,2}$ – grammar parameter, which means the number;

$k2_i = \overline{1,4}$ – grammar parameter, which means the genus;

$t_i = \overline{0,3}$ – grammar parameter, which means the time;

$h_i = \overline{1,3}$ – grammar parameter, which means the mode;

$l_i = \overline{1,3}$ – grammar parameter, which means the person;

$ch_i = \overline{1,5}$ – parameter, that is responsible for syntactical role (subject, predicate, object, definition and circumstance).

So, all characteristics of the world can be represented as unidimensional massif.

The author formulated special rules for identification context units according to the rules for creating different collocations in flexional natural languages, examples of what were depicted above. It was developed 32 rules with additions for punctuation symbols in complex sentences and for considering homogeneous parts of the sentence.

The most applicable twelve rules are represented below.

1. If the first world is adjective, numeral, pronoun or participle and the part of speech for second word is noun, their characteristics of case, number and genus are similar, the words are made collocation. For example, “*mathematical modelling*”, “*computer modelling*”, “*three pets*”, “*her name*”, “*designed room*”.

$$\begin{aligned} &if\big((cm(S_i) = 2)\, and\, (cm(S_{i+1}) = 1)\big) and \big(g(S_i) = g(S_{i+1})\big) \\ &and \big(n(S_i) = n(S_{i+1})\big) and \big(k2(S_i) = k2(S_{i+1})\big) \\ &then\big(S_j = S_i \cup S_{i+1}\big) \end{aligned}.$$

2. If the first world is noun and second word is noun of personal name too, their characteristics of case and number are similar, the words are made collocation. For example, “*Dnipro river*”.

$$\begin{aligned} &if\big((cm(S_i) = 1)\, and\, (cm(S_{i+1}) = 1)\big) and \big(g(S_i) = g(S_{i+1})\big) \\ &and \big(n(S_i) = n(S_{i+1})\big)\ then\big(S_j = S_i \cup S_{i+1}\big) \end{aligned}.$$

3. If the first world is verb and the second word is noun in genitive case, the words are made collocation. For instance, “*read book*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=5)\,\mathit{and}\,(cm(S_{i+1})=1)\big)\mathit{and}\\ \big((g(S_{i+1})=2)\vee(g(S_{i+1})=4)\big)\vee\big(g(S_{i+1})\neq 1\big).\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\big)\end{gathered}$$

4. If the first world is verb, second word is preposition and third word is noun in subjective case the words first and third are made collocation. For example, “*created for children*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=5)\,\mathit{and}\,(cm(S_{i+1})=9)\ \mathit{and}\,(cm(S_{i+2})=1)\big)\mathit{and}\\ \big(g(S_{i+2})\neq 1\big)\mathit{then}\ \big(S_j=S_i\cup S_{i+1}\cup S_{i+2}\big)\end{gathered}.$$

5. If the first world is verb and second word is pronoun not in subjective case, the words are made collocation. For instance, “*integrated scheme*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=1)\,\mathit{and}\,(cm(S_{i+1})=1)\big)\mathit{and}\big(g(S_i)=g(S_{i+1})\big)\\ \mathit{and}\big(n(S_i)=n(S_{i+1})\big)\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\big)\end{gathered}.$$

6. If the first world is participle II and second word is noun in genitive case, the words are made collocation. For example, “*using knowledge*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=7)\,\mathit{and}\,(cm(S_{i+1})=1)\big)\mathit{and}\\ \big((g(S_{i+1})=2)\vee(g(S_{i+1})=4)\big)\vee\big(g(S_{i+1})\neq 1\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\big)\end{gathered}.$$

7. If the first world is noun and second word is noun in genitive case, the words are made collocation. For example, “*modelling process*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=1)\,\mathit{and}\,(cm(S_{i+1})=1)\big)\mathit{and}\\ \big((g(S_{i+1})=2)\vee(g(S_{i+1})=4)\big)\vee\big(g(S_{i+1})\neq 1\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\big)\end{gathered}.$$

8. If the first world is noun, the second word is preposition and the third word is noun in genitive case, their characteristics of degree and number are similar, the words first and third are made collocation. For example, “*book children*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=1)\,\mathit{and}\,(cm(S_{i+1})=9)\,\mathit{and}\,(cm(S_{i+2})=1)\big)\mathit{and}\\ \big((g(S_{i+2})=2)\vee(g(S_{i+2})=4)\big)\vee\big(g(S_{i+2})\neq 1\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\cup S_{i+2}\big)\end{gathered}.$$

9. If the first world is noun and second word is verb infinitive, the words are made collocation. For example, “*indicate person*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=1)\,\mathit{and}\,(cm(S_{i+1})=5)\big)\mathit{and}\big(h(S_{i+1})=0\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\big)\end{gathered}.$$

10. If the first world is participle I and second word is adverb, the words are made collocation. For example, “*quickly integrated* ”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=6)\,\mathit{and}\,(cm(S_{i+1})=8)\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\big)\end{gathered}.$$

11. If the first world is numeral, second word is preposition and third word is pronoun in genitive case, the words first and third are made collocation. For example, “*three us*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=3)\,\mathit{and}\,(cm(S_{i+1})=9)\,\mathit{and}\,(cm(S_{i+2})=4)\big)\mathit{and}\\ \big((g(S_{i+2})=2)\vee(g(S_{i+2})=4)\big)\vee\big(g(S_{i+2})\neq 1\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\cup S_{i+2}\big)\end{gathered}.$$

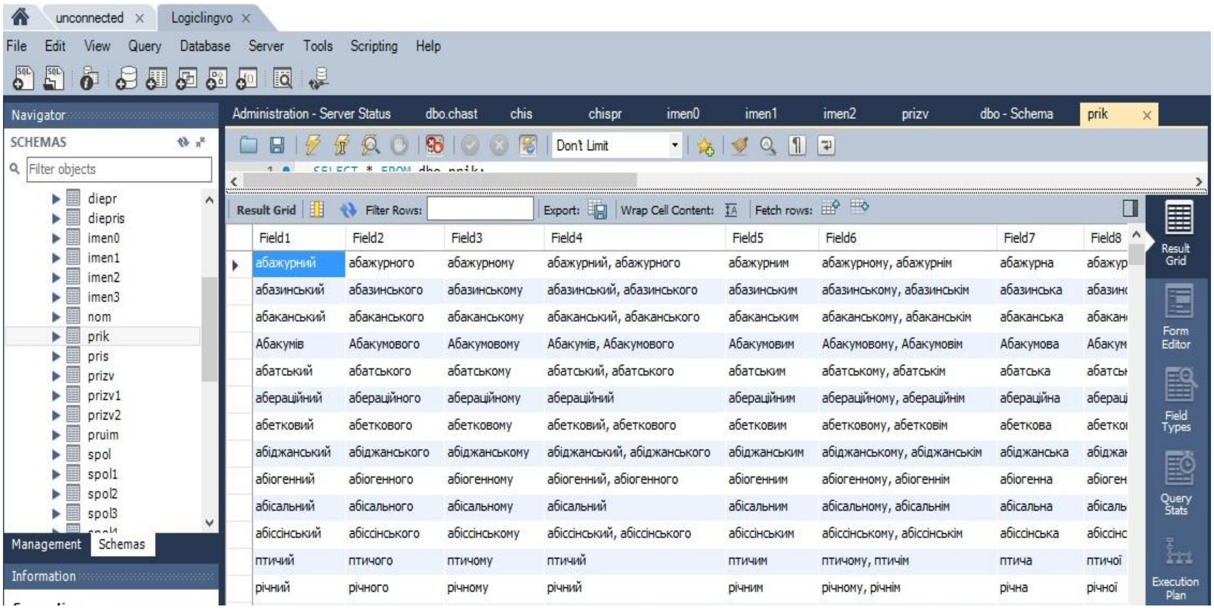
12.If the first world is numeral, second word is preposition and third word is pronoun in genitive case, the words first and third are made collocation.
For example, “*three us*”.

$$\begin{gathered}\mathit{if}\big((cm(S_i)=3)\,\mathit{and}\,(cm(S_{i+1})=9)\,\mathit{and}\,(cm(S_{i+2})=4)\big)\mathit{and}\\ \big((g(S_{i+2})=2)\vee(g(S_{i+2})=4)\big)\vee\big(g(S_{i+2})\neq 1\big)\\ \mathit{then}\big(S_j=S_i\cup S_{i+1}\cup S_{i+2}\big)\end{gathered}.$$

Created rules give us opportunity for context units identification, in spite of the words' order in natural language sentence.

3. Experiment

Using developed rules and according finding regularity it had been possible to create a system for context units identification. This system is based on usage of thesaurus. It is a tables of different grammar forms of words, where every column is responsible for particular grammar characteristics. On the Figure 9 we can see the table for adjective.



```
1 SELECT * FROM dbo.prik;
```

Field1	Field2	Field3	Field4	Field5	Field6	Field7	Field8
абажурний	абажурного	абажурному	абажурний, абажурного	абажурним	абажурному, абажурнім	абажурна	абажур
абазинський	абазинського	абазинському	абазинський, абазинського	абазинським	абазинському, абазинськім	абазинська	абазин
абаканський	абаканського	абаканському	абаканський, абаканського	абаканським	абаканському, абаканськім	абаканська	абакан
Абакумів	Абакумового	Абакумовому	Абакумів, Абакумового	Абакумовим	Абакумовому, Абакумовім	Абакумова	Абакум
абатський	абатського	абатському	абатський, абатського	абатським	абатському, абатськім	абатська	абатсы
абераційний	абераційного	абераційному	абераційний	абераційним	абераційному, абераційнім	абераційна	абераці
абетковий	абеткового	абетковому	абетковий, абеткового	абетковим	абетковому, абетковім	абеткова	абетко
абіджанський	абіджанського	абіджанському	абіджанський, абіджанського	абіджанським	абіджанському, абіджанськім	абіджанська	абіджан
абіогенний	абіогенного	абіогенному	абіогенний, абіогенного	абіогенним	абіогенному, абіогеннім	абіогенна	абіоген
абісальний	абісального	абісальному	абісальний	абісальним	абісальному, абісальнім	абісальна	абісаль
абіссінський	абіссінського	абіссінському	абіссінський, абіссінського	абіссінським	абіссінському, абіссінськім	абіссінська	абіссінс
птичий	птичого	птичому	птичий	птичим	птичому, птичім	птича	птичої
річний	річного	річному	річний	річним	річному, річнім	річна	річної

Figure 9: Example of table for adjectives

For the complex Ukrainian language sentence [10]:

"Шляхом зведення до двох різних типів систем сингулярних інтегральних рівнянь проведено чисельне дослідження задачі математичної фізики про дію стаціонарних хвиль плоскої деформації на нерухоме включення з довільним контуром, що інтегрований у нескінчене ізотропне середовище"

the system creates such context units (Figure 10).

It was made research for detection limitations and lacks of the system. For instance, identification of context units for natural language sentences, that consist unknown for database worlds or mathematical values, are not completely correct. This fact is demonstrated on Figure 11 [11]. On Figure 11 we see the words "Мак-Еліса" and incorrect word "шифрованих". The system did not recognize these words, and there are no collocations with them. So it is needed additional conditions for detection and correction mistakes in words of the natural language sentences.

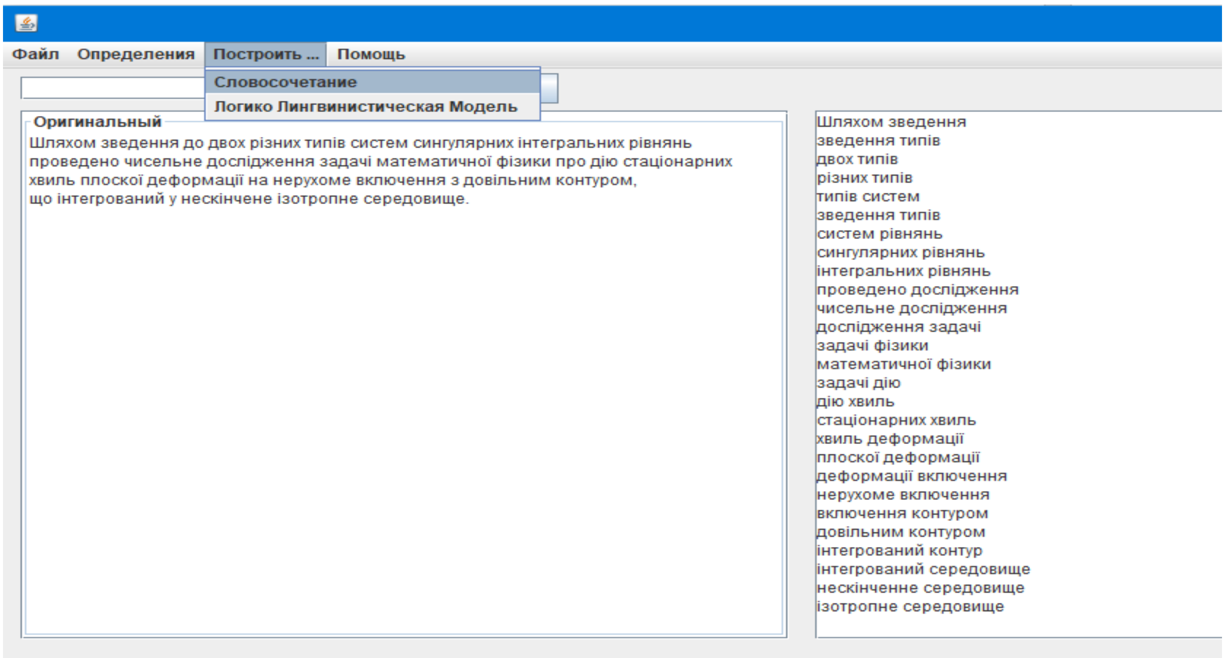
4. Conclusions

In grammatical terms, the connectivity of text is determined by the harmonization laws, rules of statement construction using morphological and syntactic means of language. In pragmatic terms, connectivity is induced by the general communicative function of the text, it is realized in subjective text organization, the system of spatial and temporal characteristics that permeate the text from beginning to end.

For this reason, it is essential to use special grammar, syntactical and semantic formal rules for context units identification.

Proposed in this article rules solve a problem of automatic identification of context collocations into the natural language sentence. For automatic generation of those colocations author suggested to use components of logic and linguistic models.

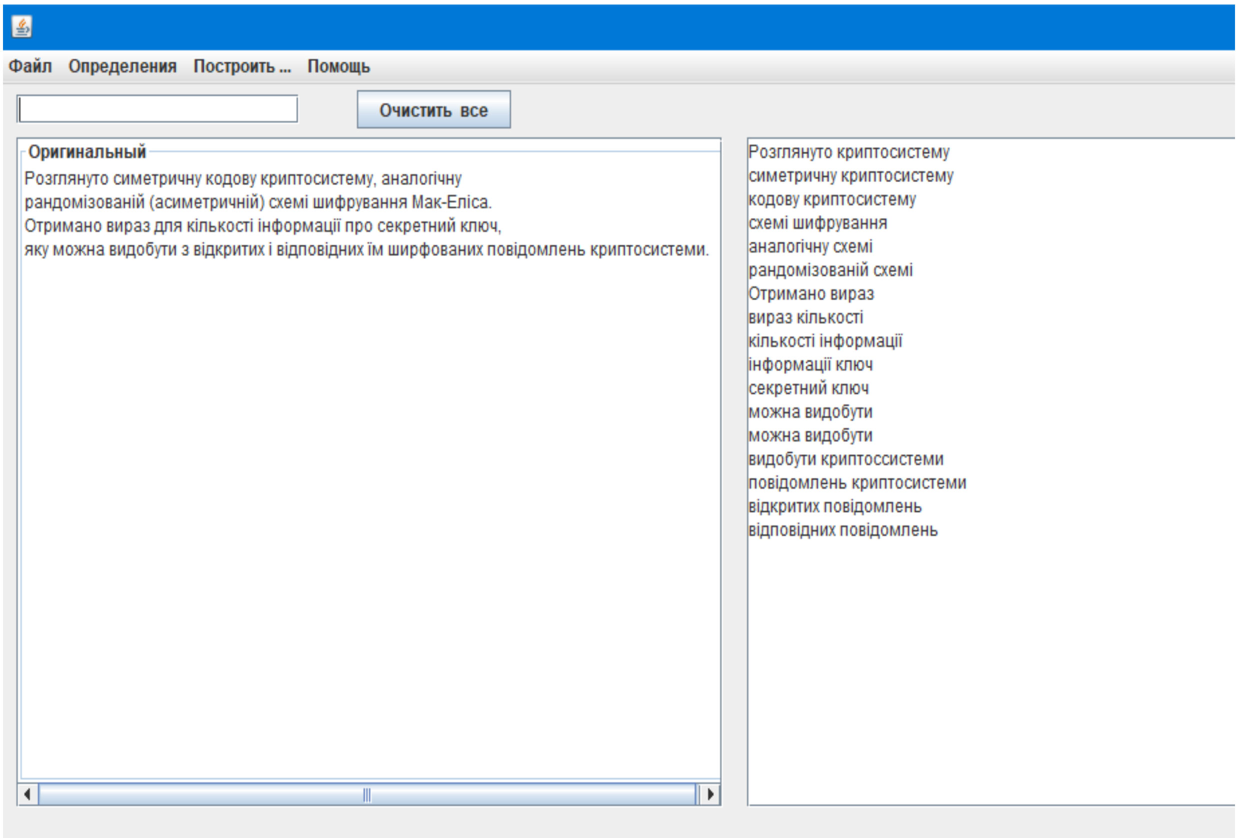
This task plays a significant part in linguistic analysis of electronic textual documents. So far as the main step in algorithm of construction of meaningful model of text is a synthesis of logic and linguistic models, based on rules of construction and searching for elementary relations.



Шляхом зведення до двох різних типів систем сингулярних інтегральних рівнянь
проведено чисельне дослідження задачі математичної фізики про дію стаціонарних
хвиль плоскої деформації на нерухоме включення з довільним контуром,
що інтегрований у нескінчене ізотропне середовище.

Шляхом зведення
зведення типів
двох типів
різних типів
типів систем
зведення типів
систем рівнянь
сингулярних рівнянь
інтегральних рівнянь
проведено дослідження
чисельне дослідження
дослідження задачі
задачі фізики
математичної фізики
задачі дію
дію хвиль
стаціонарних хвиль
хвиль деформації
плоскої деформації
деформації включення
нерухоме включення
включення контуром
довільним контуром
інтегрований контур
інтегрований середовище
нескінчене середовище
ізотропне середовище

Figure 10: Results of automatic identification of context units



Розглянуто симетричну кодову криптосистему, аналогічну
рандомізованій (асиметричній) схемі шифрування Мак-Еліса.
Отримано вираз для кількості інформації про секретний ключ,
яку можна видобути з відкритих і відповідних їм ширфованих повідомлень криптосистеми.

Розглянуто криптосистему
симетричну криптосистему
кодову криптосистему
схемі шифрування
аналогічну схемі
рандомізованій схемі
Отримано вираз
вираз кількості
кількості інформації
інформації ключ
секретний ключ
можна видобути
можна видобути
видобути криптоссистеми
повідомлень криптосистеми
відкритих повідомлень
відповідних повідомлень

Figure 11: Results of not completely full list of collocations

The above relations have identical content and conclusively interpret natural language sentences of an arbitrary structure. Logic and linguistic model of text document is a kind of pattern, which an arbitrary text document can be reduced to. Such models can be intellectual tool for searching information, word's definition, abstracting and translation.

5. References

[1] V. Evans, Lexical concepts, cognitive models and meaning-construction, in: Cognitive Linguistics, Edinburg university press Publ. Vol. 17, 2006, pp. 73-107.

[2] W. Che, Y. Zhang, Deep learning in lexical analysis and parsing, Springer Nature Singapure Pte Ltd., ch.4, in: Deep learning in Natural Language Processing, 2018. http:// doi.org/10.1007/978-981-10-5209-5_4.

[3] Y. Wilks, An Intelligent Analyzer and Understander of English, in: Communications of the ACM, Vol. 18 (5), 1975, pp. 264-274. https://doi.org/10.1145/360762.360770.

[4] Y. Zhang Discriminative syntax-based word ordering for text generation, in: Computational linguistics, Vol.41, 2015, pp. 503-538.

[5] D. Chen, C. D. Manning, A fast and accurate dependency parser using neural networks, in: *p*roceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 740–750.

[6] A. Vavilenkova, Basic principles of the synthesis of logical-linguistic models, in: Cybernetics and systems analysis, Vol. 51(5), 2015, pp. 826-834, http:// doi.org/10.1007/s10559-015-9776-z.

[7] A. Freitas, J.C. Pereira, E. Curry and P. Buitelaar, A Distributional Semantic Approach for selective Reasoninig on Commonsense Graph Knowledge bases, in: proceedings of the 19th International Conference on Applications of Natural Language to Information Systems, NLDB, 2014, Montpellier, France, pp.21-32.

[8] A. Vavilenkova, Modelling of the context links between the natural language sentences, in: Proceedings of the 9th International Scientific and Practical Conference "Information Control Systems & Technologies" (ICST2020), 2020, pp. 282-293.

[9] J. H. Martin A corpus-based analysis of context effects on metaphor comprehension, in: Corpus-Based Approaches to Metaphor and Metonymy, Berlin, 2006, pp. 214-236, https://doi.org/10.1515/ 9783110199895.214.

[10] B.E. Panchenko, Y.D. Kovalev and I.N. Saiko, Numerical Analysis of Systems of Singular Integral Equations of the First Kind with an Indefinable Index in the Problem of Diffraction of Plane Waves on a Rigid Inclusion, in: *Cybernetics and System Analysis, Vol.* 56, 2020, pp. 521–533. https://doi.org/10.1007/s10559-020-00268-z.

[11] S.V. Mitin, Amount of Key Information Contained in Plain and Encrypted Text Sets of the Symmetric Randomized McEliece Cryptosystem, in: *Cybernetics and System Analysis, Vol.* 56, 2020, pp. 726–730. https://doi.org/10.1007/s10559-020-00288-9

[12] F. H. George The foundations of cybernetics, Gordon and breach science publishers Ltd., U.K., 1977.

[13] Stefan Th. Gries, A. Stefanovich (eds.), Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis, in: Trends in Linguistics, Berlin/New York, Mouton de Gruyter, 2006.

[14] N. Lukashevich, Thezauruses into the tasks of information searching, SPB, Published by Vilyams, 2011.

[15] A. Bashmakov, I. Bashmakov Intellectual information technologies, M., MGTU Publish, 2005.

[16] M. Plusch, N. Grupas Ukrainian language, Kiyv, Published by Radyanska school, 1990.

[17] A. Vavilenkova, Analysis and synthesis of logic and linguistic models, TOV "SIK GROUP Ukraine", 2017.

[18] Progaonline.com, 2020, URL: https://progaonline.com/syntax/result/ cc86fc72309e0d21c2590c073df987d3.

[19] Automatic text processing, 2020, URL: http://aot.ru/.

[20] Erg.delpfi-in, 2020, URL: http://erg.delph-in.net/logon.

[21] Link Grammar, 2020, URL: http://www.link.cs.cmu.edu/link.