

**TURLI SO‘Z TURKUMLARI ORASIDAGI OMONIMIYANI ANIQLOVCHI
MATEMATIK MODELLAR****Axmedova X.I**

Alisher Navoiy nomidagi O‘zbek tili va adabiyoti universiteti

Tayanch doktoranti

<https://doi.org/10.5281/zenodo.7238546>

Annotatsiya. Ushbu maqolada o‘zbek tilidagi omonim so‘zlarni semantik tahlil qilish masalasi ko‘rib chiqiladi. Omonim so‘zlar ikkita: bir so‘z turkumi va turli so‘z turkumlari orasidagi omonimlik guruhlarga ajratilgan holda matematik model va algoritmlar ishlab chiqilgan. Omonimlar so‘zlarning turli turkumlarga mansub bo‘lish ierarxiyasi shakllantirilgan bo‘lib, maqolada ravish ∨ fe‘l, ot ∨ ravish, sifat ∨ son, ot ∨ son, ot ∨ son ∨ fe‘l kabi so‘z turkumlariga mansub omonimlarni aniqlash usullari keltirilgan. Ushbu algoritmlardan o‘zbek tili semantik analizatori dasturiy ta‘minotini yaratishda foydalanish mumkin.

Kalit so‘zlar: matematik model, taqlid so‘z, ot, fe‘l, ravish, son, omonim, ketma-ketlik, to‘plam, chekli to‘plam, semantik analizator.

**MATHEMATICAL MODELS DETERMINING HOMONY BETWEEN
DIFFERENT WORD GROUPS**

Abstract. This article discusses the issue of semantic analysis of homonymous words in Uzbek language. Mathematical models and algorithms have been developed by dividing homonymous words into homonymous groups between a single word group and various word groups. A hierarchy of homonyms belonging to different categories of words is formed. The article presents methods for identifying homonyms belonging to such word groups as adverb ∨ verb, noun ∨ adverb, adjective ∨ numeral, noun ∨ numeral, noun ∨ numeral ∨ verb.

Keywords: mathematical model, imitation word, noun, verb, adverb, numeral, homonym, sequence, set, finite set, semantic analyzer.

**МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОПРЕДЕЛЕНИЯ ОМОНИИ МЕЖДУ
РАЗНЫМИ СЛОВОГРУППАМИ**

Аннотация. В данной статье рассматривается вопрос семантического анализа омонимичных слов в узбекском языке. Математические модели и алгоритмы были разработаны путем деления омонимичных слов на омонимичные группы между одной группой слов и различными группами слов. Формируется иерархия омонимов, принадлежащих к разным категориям слов. В статье представлены методы выявления омонимов, принадлежащих к таким группам слов, как наречие ∨ глагол, существительное ∨ наречие, прилагательное ∨ числительное, существительное ∨ числительное, существительное ∨ числительное ∨ глагол.

Ключевые слова: математическая модель, имитационное слово, существительное, глагол, наречие, числительное, омоним, последовательность, множество, конечное множество, семантический анализатор.

KIRISH

Kompyuter lingvistikasi sohasining muhim masalalaridan biri matnlarni semantik tahlil qilish hisoblanadi. Matnlarni semantik tahlil qilish bir qancha elementlarga ega:

- giponimiya;
- omonimiya;
- polisemiya;
- polifunksionallik;
- sinonimiya;
- antonimiya.

Semantik analizning muhim elementlaridan biri omonimiya hosidasidir. Omonimiyani aniqlash turli tabiiy tillarda turlicha talqin qilinadi. Omonim soʻzlarni farqlashda ularni ikkita guruhga ajratamiz: bir soʻz turkumi va turli soʻz turkumlari orasidagi omonimlik [1]. Sh. Gulyamova tomonidan taklif etilgan turli soʻz turkumlari orasidagi omonimlikni farqlovchi lingvistik modellarga [2] tayangan. holda matematik model va algoritmlar ishlab chiqamiz.

MAVZUGA OID ADABIYOTLAR TAHLILI

Jahon kompyuter lingvistikasida gaplarni semantik tahlil qilishda asosan 3 ta usuldan foydalanilgan:

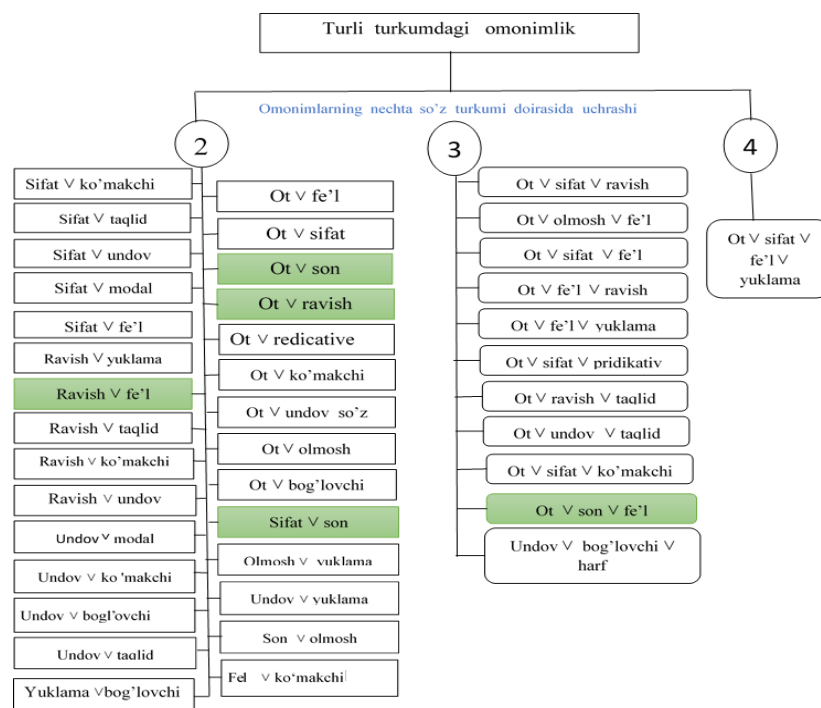
- Qoidalarga asoslangan usul;
- Statistik maʼlumotlarga asoslangan usul;
- Mashinali oʻrganishga asoslangan usul.

Ushbu usullar turli tillarda turlicha foydalanilgan. Masalan, Rus tilshunosligida omonimlarni oʻrganishga bagʻishlangan tadqiqotlar juda koʻp. Omonimiya hodisasi A.A.Poroxin [2013], D.N.Gomon [2004], D.A. Mixaylovna [2015], P. Boris Kobritsov, Olga N. Lashevskaja, Olga Ju. Shemanaeva [2011], A.I. Bolshakova [2003], B. A. Bobnev (2010), S.V.Rysakov, E.S.Klyshinskiy [2015], A.V. Gashkov [2012] va boshqalarning ishlarida maxsus tadqiq etilgan. Rus tilida omonim soʻzlarni farqlashda asosan qoidalarga asoslangan usul va statistik usullardan foydalanilgan.

Omonimiya Baltabayeva J.K. va Sulaymonova J.N (qozoq tili) [2019], Ch.A.Davlyatova (tojik tili) [2017], V.V. Kukanova(boshqird tili) [2014], H.Geydarova(ozarbayjon tili) [2017] V.V. Kukanova(qalmoq tili)[2011] kabi turkolog olimlarning tadqiq mavzusi boʻldi. Bu turkolog olimlar A.I.Smirniskiy, V.V.Vinogradov, O.S.Axmanova va boshqa tilshunos olimlar ishlab chiqqan omonimiya nazariyasiga tayanadilar. Turkolog olimlarning izlanishlari natijasi shuni koʻrsatadiki omonimiyani aniqlash yuqorida keltirilgan metodlar muhim ahamiyatga ega. Ushbu maqolada oʻzbek tilida turli grammatik jihatdan oʻxshash boʻlmagan soʻz turkumlari doirasidagi omonimiyani farqlashda qoidalarga asoslangan usullardan foydalanish qulayligi asoslanadi va shunday tipdagi omonimlarni farqlovchi matematik modellar tavsita etiladi.

TADQIQOT METODOLOGIYASI

Turli soʻz turkumlariga doir omonim soʻzlar orasidan toʻp soʻzi misolida morfologik tahlilni amalga oshiramiz. Matndagi omonim soʻzlar farqlovchi lingvistik modellar chet ellik olimlar tomonidan tahlil qilingan [4-8]. Oʻzbek tilidagi omonim soʻzlarni turkumlar doirasida guruhlariga ajratib farqlashni tavsiya qilamiz. Bir soʻz turkumi doirasidagi omonim soʻzlarni staxostik va semantik konkardans usuli yordamida, turli soʻz turkumlari doirasidagi omonim soʻzlarni qoidalarga asoslangan usullar yordamida farqlashni tavsiya qilamiz.



1-rasm. Turkumlar doirasidagi omonim so`zlar ierarxiyasi.

O`zbek tilidagi omonim so`zlarning turli turkumlar (2, 3 va 4 ta)ga mansub bo`lish ierarxiyasi 1-rasmda keltirilgan bo`lib, ushbu maqolada quyidagi turkumlar tahlil qilinadi:

- ravish √ fe`l;
- ot √ son;
- ot √ ravish;
- sifat √ son;
- ot √ son √ fe`l;

—

I. Ot √ taqlid so`z turkumlari doirasidagi omonimiyani farqlovchi modellar

O`zbek tilidagi matnlarda to`p so`zi 4 xil shaklda ot so`z turkumi sifatida yoki 1 ta taqlid so`z vazifasida uchraydi. To`p so`zi ot so`z turkumiga oid bo`lgan hollarda uni bir so`z turkumi doirasidagi omonimlarni farqlash usuli yordamida farqlanadi[3]. Buni inobatga olgan holda ot so`z turkumi √ taqlid so`zni farqlovchi leksik-morfologik omillarni qaraymiz:

- 1) to`p: Koptok, futbol koptogi. (ot);
- 2) to`p: Snaryad otuvchi uzun stvolli harbiy qurol, zambarak. (ot);
- 3) to`p: So`zlash. Peshin.(ot);
- 4) to`p: O`ram, bog`lam, guruh. (ot);
- 5) to`p: Yumshoq jismning yerga urilishidan chiqadigan tovushni bildiradi. (taqlid so`z);

To`p so`zi taqlid so`z bo`lib kelganida unga qo`shilishi mumkin bo`lgan qo`shimchalarni ham, brikuvchi so`zlarni ham aniqlab olish talab qilinadi. Yaratilgan lingvistik modellar [2]ga tayanib quyidagi matematik mulohazani keltiramiz:

$$H^{imit} = \begin{cases} H + aff_i^{CS} \oplus V_j^{imit} \\ H \oplus \downarrow N_i^{imit} + aff_i^N \oplus V_j^{imit} \\ H \oplus V_k^{q2} \end{cases} \quad (1)$$

Bu yerda, $[[aff]] _i^{CS}$ - kelishik qo'shimchalar to'plamining qism to'plami, ya'ni

Tushum kelishigi –ni: to'pni bas qil;

Jo'nalish kelishigi –ga: to'pga uyg'on/kel;

Chiqish kelishigi –dan: to'pdan qo'rq/uyg'on, demak taqlid so'zlar otlashish xususiyatiga ega ekan.

V_j^{imit} - taqlid so'zlarga brikuvchi fe'l so'z turkumidagi so'zlardan iborat chekli to'plam, masalan:

$V_j^{imit} = \{bas\ qilmoq, uyg'onmoq, kelmoq, \dots\}, j=1 \dots n$

Taqlid so'z otlashganda otga xos grammatik shaklni qabul qilishi so'zni ot-omonimga o'xshab qolishiga va ayrim birikuvchilar bir xil bo'lishiga olib keladi. Shuni hisobga olgan holda

N_l^{imit} - taqlid so'z bilan brika oladigan ot so'z turkumiga oid so'zlardan iborat chekli to'plam hosil qilinadi. $[[aff]] _i^N$ - ot so'z turkumiga qo'shilishi mumkin bo'lgan qo'shimchalar:

Tushum kelishigi –ni: to'p degan ovozni bas qil;

Jo'nalish kelishigi –ga: to'p ovoz+i/lari/miz/...+ga uyg'on/kel;

Chiqish kelishigi –dan: to'p ovoz+i/lari/miz/...+dan qo'rq.

V_k^{q2} -qo'shma fe'lning 2-komponenti, ya'ni taqlid so'zga qo'shib qo'shma fe'l hosil qiladi:

$V_k^{q2} = \{etmoq, demoq, qilmoq, \dots\}$

To'p qilmoq, to'p demoq, ...

Yuqorida keltirilgan mulohazalar orqali ot \vee taqlid so'z turkumlari orasidagi omonimlikni farqlash mumkin.

II. Xuddi shunday ravish \vee fe'l so'z turkumlari orasidagi omonimlikni farqlovchi modellarni ham keltiramiz:

Oz: Kam, miqdori nisbatan ko'p bo'lmagan. (ravish)

Oz: Oriqlamoq, etidan yo'qotmoq. (fe'l)

Oz: Noxush bo'lmoq, kuchsizlanmoq, holsizlanmoq. (fe'l)

Oz: Adashmoq, to'g'ri yo'ldan chetga chiqmoq. (fe'l)

Oz so'zi 3 ta o'rinda fe'l va bitta ravish so'z turkumi vazifasida uchraydi. Bu tipdagi omonim so'zlarni farqlash uchun ularni leksik va morfologik tahlil qilishimiz kerak. Fe'l so'z turkumining leksik-morfologik xususiyatlarini [3] inobatga olib, ravish so'z turkumini leksik-morfologik tahlil qilamiz.

oz+(-gina)+odam/mehnat/mashaqqat/bola/vaqt;

oz+gapir/so'zla/kul/o'yna/ishla/tishla/qol/kut/sabr qil/sarfla;

oz ravishi otlashish xususiyatiga ega bo'lib, oz fe'lida bunday hodisa kuzatilmaydi.

Ravishga –i egalik qo'shimchasi qo'shilsa u old tomonidan ot so'z turkumiga oid so'z bilan birikadi va unga qaratqich kelishigi qo'shimchasi qo'shiladi, ya'ni

gapning ozi yaxshi, ishning ko'pi;

mevaning ozini olib kelmoq;

Bu qoidalarni inobatga olgan holda quyidagi modelni keltiramiz

$$H^{adv} = \begin{cases} H + \downarrow aff_i^{adv} \oplus N, & i=1..n, aff_i^{adv} \in aff^N \\ H + \downarrow aff_i^{adv} \oplus V, & i=1..n, aff_i^{adv} \in aff^N \\ N+ning \oplus H+i+ \downarrow ni \oplus V \end{cases} \quad (2)$$

Bu yerda $[[aff]] _i^{adv}$ -ravishga qo‘shilishi mumkin bo‘lgan qo‘shimchalar, yani kelishik, egalik, o‘tning kichraytirish erkalash shaklini ifodalovchi qo‘shimchalar, sifat darajasi va ularning kombinatsiyalaridan iborat qo‘shimchalarning chekli to‘plam: $[[aff]] _i^{adv} = \{-gina, -roq, -roqqina, -si, -ginasi,$

$-ning, -ginasini, -ni, \dots\}$, N -ot, V -fe‘l

III. Son turkumidagi so‘zning ot, sifat, fe‘l so‘z turkumiga oid so‘zlar bilan omonimlik hosil qilishi mumkinligi kuzatildi .

Yetti, yigirma so‘zlari shular qatoridandir. Yetti so‘zi son va ot so‘z turkumlari orasida omonimlik hosil qiladi.

Yetti: Olti bilan sakkiz orasidagi butun sonning nomi. (son)

Yetti: Kishi vafotining to‘rtinchi kuni o‘tkaziladigan Marosim.(son)

Yetti so‘zini leksik- morfologik tahlili:

Yetti (son)+ kishi /so‘m /metr /kun /mahalla /xufton /iqlim /og‘ayni /qaroqchi /pusht (ot) ;

Yetti (son)+ o‘lcha/bukil (fe‘l);

Yetti (ot) + $\{-ni/-ga/-da/-dan\} + \{o‘tkaz /bo‘l /ko‘rish /gaplash /xayrlash /bor\}$ (fe‘l) // + $\{-ning\} + \{kelishi/bo‘lishi\}$ (harakat nomi).

Son \vee ot so‘z turkumlari orasidagi omonimlikni hosil qiluvchi so‘zlarning barchasi shu kabi tahlil qilinadi.

$$H_{num,N} = \begin{cases} \downarrow W_i^{dec} \oplus H^{num} + \downarrow aff_i^{ord} \oplus \downarrow (W_k^N + \downarrow aff_i^{ord}) \oplus N_j^c \\ H^{num} \oplus V^{num} \\ H^N + aff^{CS} \oplus V \end{cases} \quad (3)$$

Bu yerda H^{num} - son so‘z turkumiga oid omonim so‘z, aff^{ord} (ordinal number)- tartib sonlarni bildiruvchi qo‘shimchalar $aff^{ord} = -nchi, -inchi, -ta, -tacha, \dots$; W^{dec} -o‘nliklar, yuz, ming va boshqa sonlarning nomlaridan iborat chekli to‘plam, $W^{dec} = o‘n, yigirma, o‘ttiz, \dots, yuz, ming, million, milliard, trillion, \dots, dodekalon, i=1..23$; W^kN - raqam so‘zlar, $W^kN = bir, ikki, \dots, to‘qqiz, k=1..9$, N^j -sanaladigan otlar to‘plami, V^{num} -son omonimga birikuvchi fe‘llar to‘plami, H^N -son omonimning ot so‘z turkumi vazifasida kelgan ko‘rinishi, aff^{CS} - kelishik qo‘shimchalari to‘plami.

IV. Quyida son \vee sifat so‘z turkumlari orasidagi omonimlikni farqlovchi modellar ko‘rib chiqamiz.

1) Ellik : Beshinchi o‘nlikning nomi.(son)

2) Ellik : Barmoqning kengligiga teng, serbar.(sifat)

Son va sifat so‘z turkumlarining leksik- morfologik tahlili keltiramiz.

- {Ellik}(son) ⊕ {bosh, nafar, dona, marta, karra, bir, ikki,...}(ot);
- {yuz, ming, million,...} ⊕ {Ellik}(son)+ta/-tacha/-inchi/... ⊕ (ot).

Ellik so‘zi sifat so‘z turkumi vazifasida quyidagicha uchraydi.

1. Ellik gaz gazlama sotib olindi
 2. Ellik yog‘di.
- Ellik (sifat) +{yog‘} (fe‘l);
 - Ellik (sifat) +{chang} (ot).

Ushbu leksik-morfologik xususiyatlarni inobatga olgan holda quyidagi matematik mulohazani taklif qilamiz.

$$H = \begin{cases} H^{num} + \downarrow aff_l^{ord} \oplus \downarrow (W_k^N + \downarrow aff_l^{ord}) \oplus N_j^c \\ H^{adj} \oplus V \\ H^{adj} \oplus N \end{cases} \quad (4)$$

V. O‘zbek tilida shunday omonimlar borki, kontekstlarda 3 ta so‘z turkumi doirasida uchraydi.

- 1) Uch: Biror narsaning boshlanish yoki tugash qismi. (ot)
- 2) Uch: Ikki bilan to‘rt orasidagi butun son.(son)
- 3) Uch: 1. Havo bo‘ylab harakatlanmoq. (fe‘l)
2. Biror qirrasini cho‘rt uzilib tushmoq.
3. Beixtiyor berilmoq, laqillamoq

Uch so‘zi quyidagicha leksik-morfologik va sintaktik tahlil qilinadi:

- uch (ot)+-i/-iga/-ni/-da ⊕ b\or/ulamq/boshlamq/turmoq;
- uch (son)+-inchi/-nchi/-ta/-tacha/-tadan/-larcha/ ⊕ oy nomlari/dars/uy/...;
- uch (son)+-ov/-ovlon/... ⊕ gaplashmoq/ko‘rishmoq/kelmoq/ketmoq/;
- uch (son)+dan ⊕ bir/ikki/uch/to‘rt/...;
- Balandga/osmonga ⊕ uch (fe‘l);
- Toshkentdan/shahardan... ⊕ uch (fe‘l);
- Uch+-ib ⊕ kel/ket;
- Uch+-ar ⊕ qush/gilam;

Uch so‘zning leksik-morfologik va sintaktik tahlili asosida ot, son yoki fe‘l so‘z turkumlari orasidagi omonim so‘zlar orasidan son so‘z turkumini ajratib olish uchun quyidagi matematik mulohazani taklif qilamiz.

$$H^{num} = \begin{cases} H^{num} + \downarrow aff_l^{ord} \oplus \downarrow (W_k^N + \downarrow aff_l^{ord}) \oplus N_j^c \\ \downarrow W_i^{dec} \oplus H^{num} + \downarrow aff_l^{ord} \oplus N_j^c \\ H^{num} + aff_j^{agg} \oplus V \\ H^{num} + dan \oplus W_n^{num} \end{cases} \quad (5)$$

Bu yerda [[aff]]_j^agg-jamlovchi qo‘shimchalar to‘plami, [[aff]]_j^agg={-ov,-ovlon,-ala,-ovi,-oviga,-ovloni,...}, W_n^num-sonlarning nomlaridan iborat so‘zlar to‘plami, W_n^num={bir,ikki,uch,...}

Ot va fe'l orasidagi omonimlikni farqlashda [1] taklif qilingan matematik mulohazalardan foydalaniladi.

TAHLIL VA NATIJALAR

Ishlab chiqilgan matematik modellar asosida semantic analizator axborot tizimi uchun algoritmlar ishlab chiqiladi. Ishlab chiqilgan algoritmlar asosida dasturlash tillari yordamida funksiya va metodlar yoziladi. Matematik modellarning ahamiyatini aniqlash maqsadida yozilgan funksiyalar shuni ko'rsatadiki, masalan ot va taqlid so'z turkumidagi omonim so'zni faqat so'z turkumini 80-82 % aniqlik bilan farqlaydi. Qolgan 18-20 % aniqlikni esa statistik va mashinali o'rganishga asoslangan usullar yordamida aniqlani. Boshqa matematik modellar xuddi shunday aniqlik bilan omonimiyani farqlashga yordam beradi. Matematik modellarni to'g'ri ishlashi uchun albatta zarur so'z va qo'shimchalar guruhlari to'plamlari elementlar bilan to'ldirilgan bo'lishi talab qilinadi

XULOSA

Turli turkum doirasida kuzatiladigan omonimlarni modellashtirish ularning aynan qaysi turkumlar orasida ekanligini aniqlashdan iborat. Ism tarkibiga kiruvchi turkumlar orasida bo'lsa, ularni ko'p holda grammatik shakllar orqali farqlab bo'lmaydi. Ba'zan ism tarkibiga kirmaydigan turkum ham xuddi ismlardek uning lug'aviy va sintaktik shakllarini qabul qilib kelishi mumkin. Bu holatda sintaktik omil yetakchi o'ringa chiqib, birikuvchanlik tamoyili asosida farqlanadi va shunga muvofiq modellashtiriladi.

REFERENCES

1. Гулямова Ш.Қ., Ахмедова Х.И. Ўзбек тили семантик анализатори учун омоним сўзларлар маълумотлар базасини шакллантириш масаласи хусусида// Сўз санъати халқаро журнали. 4-жилд, 3-сон. – Тошкент, 2021.– Б. 326-334. – Doi 10.26739/2181-9297
2. Гулямова Ш.К. Семантик анализатор учун омонимликни фарқлаш омилларининг айрим масалалари хусусида // “Ўзбекистоннинг умидли ёшлари” мавзусидаги 6-сон Республика илмий талабалар, магистрлар ёш тадқиқотчилар ва мустақил изланувчилар учун онлайн конференцияс и н и н г материаллари тўплами. – Тошкент: Tadqiqot, 2021. – Б. 74-76.
3. Axmedova X.I. Mathematical models that distinguish homonymy in the framework of a word series. Electronic journal of actual problems of modern science, EDUCATION AND TRAINING OCTOBER, 2021-10/1. ISSN 2181-9750
4. Krzysztof K. Формальные и семантические сближения в сфере русско-польской межъязыковой омонимии. Studia Rossica Posnaniensia, (2021), 46(1). DOI:10.14746/strp.2021.46.1.8
5. Andrey K. Ilona K. Clustering of word contexts as a method of eliminating polysemy of words. Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, EIConRus 2019 (2019). DOI: 10.1109/EIConRus.2019.8656851
6. Jonathan P. Bona, Werner Ceusters. Mismatches between major subhierarchies and semantic tags in SNOMED CT. Journal of Biomedical Informatics (2018). DOI: 10.1016/j.jbi.2018.02.009

7. Ondřej P., Miloslav K. UlsaNA: Universal language semantic analyzer. International Conference Recent Advances in Natural Language Processing, RANLP (2019). DOI: 10.26615/978-954-452-056-4_112
8. Anastasia M., Vladimir M. Mathematical model of an ontological-semantic analyzer using basic ontological-semantic patterns. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2017). DOI: 10.1007/978-3-319-62434-1_5