

## ლინგვისტური პროცესორები: მიმოხილვითი ანალიზი

ბადრი მეფარიშვილი, თამარ მეფარიშვილი,  
გულნარა ჯანელიძე  
საქართველოს ტექნიკური უნივერსიტეტი,  
თბილისის სახელმწიფო უნივერსიტეტი

### რეზიუმე

ნაშრომი ეძღვნება ლინგვისტური პროცესორების განვითარების პრობლემებს. ჩატარებულია თანამედროვე HTML და XML ტექნოლოგიების შედარებითი ანალიზი. შემოთავაზებულია განხილული პრობლემების გადაწყვეტის ახალი მიდგომა. ტექსტური პროცესორის რეალისტური მოდელი ითვალისწინებს აგრეთვე ბუნებრივი ენის ტექსტის გაგებას კომპიუტერის მიერ, ე.ი. როგორც შემეცნების ობიექტის, ასევე მისი ამოცნობის შესაბამისი პროცედურის არსებობას. შემოთავაზებულ მიდგომაში შემეცნების ობიექტად წარმოდგენილია ხელოვნური კონსტრუქცია – შემეცნებითი სახე. დასკვნაში შემოთავაზებული იდეის რეალიზაცია წარმოდგენილია როგორც ხელოვნური ინტელექტის განვითარების სამომავლო პარადიგმა.

**საკვანძო სიტყვები:** ლინგვისტური პროცესორი. ხელოვნური ინტელექტი. ბუნებრივი ენა. სახეთა ამოცნობა. HTML და XML ტექნოლოგიები.

### 1. პრობლემის არსი და აქტუალობა

თანამედროვე საინფორმაციო-საძიებო სისტემების ფუნქციური არეალი მნიშვნელოვნად გაფართოვდა. ასეთი სისტემები შეიძლება გამოყენებულ იქნას ადამიანის საქმიანობის მრავალ სფეროში, სადაც ინფორმაციის უდიდესი ნაწილი წარმოდგენილია დოკუმენტების სახით. სამწუხაროდ, ისინი ჯერ კიდევ ვერ პასუხობენ ბუნებრივ ენაზე ფრაზების შეცნობისა და სინტაქსის მიმართ წაყენებულ მკაცრ მოთხოვნებს, შესაბამისად ვერ უზრუნველყოფენ მომხმარებელთან დიალოგის წარმართვის მაღალ ხარისხს.

პრაქტიკულად საინფორმაციო-საძიებო სისტემებში გამოიყენება შეზღუდული ბუნებრივი ენა [3]. ასე მაგალითად, Internet-ში საჭირო დოკუმენტების მოძიებისათვის ძირითადად გამოიყენება მხოლოდ გასაღებური სიტყვები. მდგომარეობას ისიც ართულებს, რომ მომხმარებლები უფრო მეტად იყენებენ მათთვის მისაღებ სხვადასხვა სპეციალურ ტერმინებს, რის გამოც სწორად წარმოიშვება გასაღებურ სიტყვებთან მათი ტერმინების შეუსაბამობის პრობლემა.

ენათმორისი კომუნიკაციების ზრდადი მოცულობები, ენობრივი ბარიერების გადალახვის აუცილებლობა სულ უფრო მეტად ითხოვს ტექსტების ავტომატური დამუშავების სერიოზულ ტექნოლოგიურ მხარდაჭერას. მოცემულ კონტექსტში მკვეთრად არადაამაკმაყოფილებელია მანქანური თარჯიმანი სისტემების დონე, რომლებიც ძირითადად სიტყვა-სიტყვით თარგმანს ახორციელებენ, რითაც შინაარსობრივი მხარე ზარალდება. გარდა ამისა, Web განიხილება როგორც პოტენციალური ცოდნის ბაზა, რომელთანაც მუშაობისათვის საჭიროა მოთხოვნებისა და ცოდნის წარმოდგენისა თუ დამუშავების სპეციალური მეთოდები, შესაბამისი ინტერფეისი. ამ

თვალსაზრისით, მიზანშეწონილად შეიძლება ჩაითვალოს ხელოვნური ინტელექტის მეთოდებისა და საშუალებების გამოყენება [5].

საინფორმაციო-საძიებო სისტემების ერთერთ საკვანძო კომპონენტად გვევლინება **ლინგვისტური პროცესორი** (ლპ), როგორც მომხმარებელსა და მონაცემთა ბაზებს შორის შუალედური რგოლი, რომლის ძირითად ამოცანას წარმოადგენს ბუნებრივ ენაზე შედგენილი წინადადების (ან მთელი ტექსტის) შინაარსის ფორმალური წარმოდგენის შესაბამისი სემანტიკური სტრუქტურების გარკვეულ კრებულად გარდაქმნა.

მიუხედავად ლინგვისტური პროცესორების შემუშავების ნახევარსაუკუნოვანი ისტორიისა და მრავალრიცხოვანი ლინგვისტური მოდელებისა, რომლებიც ძირითადად განსხვავდებიან ლინგვისტური ფენომენების (მორფოლოგიის, სინტაქსისის, სემანტიკისა და სხვ.) აღწერის თავისებურებებით, ჯერჯერობით ვერც ერთმა შემოთავაზებულმა მიდგომამ დამაკმაყოფილებელი შედეგი ვერ გამოიღო [6]. გლობალური ქსელების შექმნამ, ტექსტური რეპოზიტორიების ზრდამ ბუნებრივ ენაზე ტექსტების ავტომატური დამუშავების ტექნოლოგიურ პრობლემებს გასაკუთრებული მნიშვნელობა შესძინა.

ლინგვისტური პროცესორების განვითარების მთავარ სირთულეს წარმოადგენს აზროვნებისა და მეტყველების მოდელებს შორის პრინციპული შეუსაბამობა. ტვინი, რომელიც წარმოადგენს მოქნილი, გადაწყობადი სტრუქტურის, მრავალდონიან ნეირონულ სისტემას, მანიპულირებს დიდი განზომილების სუპერგეომეტრული სახეებით. ტვინის ნეირონულ ქსელებში არსებული ცოდნის მოდელი ისევე შეუსაბამოა ფრეიმების სახით წარმოდგენილ ხელოვნური ცოდნის თანამედროვე მოდელთან, როგორც მაღალი განზომილების სუპერგეომეტრია წრფივ ან თუნდაც ორგანზომილებიან გეომეტრიასთან. ამდენად, დღემდე არსებული ყველა მიდგომა არის ხელოვნური ინტელექტის ბუნებრივთან მხოლოდ პირველი რიგის მიახლოება [2].

ზოგადად, პრაქტიკულ ასპექტში ინტელექტუალური ლინგვისტური პროცესორები მოწოდებული არიან გადაწყვიტონ შემდეგი ამოცანები:

- მანქანური თარგმანი;
- ინტელექტუალური ძებნა (ტექსტური დოკუმენტების ძებნა მოთხოვნის შინაარსის და არა გასაღებური სიტყვების მიხედვით). მოცემულ კონტექსტში განსაკუთრებულ მნიშვნელობას იძენს ე.წ. Ad hoc query-ის რეალიზაცია დიდი განზომილებისა და რთული სტრუქტურის მონაცემთა ბაზებში;
- გაფანტული და განაწილებული ინფორმაციის შეკრება გარკვეული თემის, პრობლემის, პერსონალების მიხედვით, აგრეთვე, ინფორმაციული დაზვერვის თვალსაზრისით, ყველა შესაძლო დოსიების შედგენა და მხარდაჭერა, რომლებიც საჭიროებს მონიტორინგს, შევსებას თუ აქტუალიზაციას;
- ხლომილობებს, პერსონებს შორის კავშირების განსაზღვრა და ა.შ.;

- კორპორაციული სამსახურების ფუნქციონირების უზრუნველყოფა შემავალი ელექტრონული ფოსტის შინაარსის მიხედვით (ანალიზი, სორტირება, შემსრულებლებზე გადამისამართება, ე.წ. „სპამის“ ფილტრაცია, ტენდენციების გამოვლინება და სხვა) ავტომატური დამუშავების თვალსაზრისით;

- მარკეტინგული ინფორმაციის ძებნა და სტრუქტურირება, მონაცემთა ბაზების შედგენა მომხმარებლების, პრიორიტეტების, კონიუნქტურული მაჩვენებლების, პროფილების მიხედვით და სხვა;

- მაღალი რისკის მართვის სისტემის საიმედოობის უზრუნველყოფა, კერძოდ გადაწყვეტილების მიმღებ პირების (გმპ) ფსიქომეტრულ-ვერბალური ტესტირებისა და გადაწყვეტილებათა ადექვატურობის მონიტორინგის თვალსაზრისით;

- ექსპერტების ცოდნათა ტექსტურიდან სემანტიკურ ფორმაში კონვერტირება, ფორმალიზაცია მათი შეჯერებისა და შემდგომ ექსპერტულ სისტემებში ცოდნის ბაზების ფორმირებისა და თვითგანსწავლის მიზნით და სხვ.

ამგვარად, სოციალურ-კომუნიკაციური ამოცანების გადაწყვეტის ასპექტში, მთელი სიცხადით დადგა ახალი თაობის ინტელექტუალური ლინგვისტური პროცესორების შემუშავების აქტუალობა.

## **2. ლინგვისტური პრობლემები**

ცივილიზაციის ევოლუციურმა პროცესებმა განაპირობეს ფუნქციონალურ-კომუნიკაციური ლინგვისტიკის განვითარება. დემოგრაფიულმა დიფუზიამ გამოიწვია როგორც ლექსიკური, ისე გრამატიკური მოდელების ურთიერთგამდიდრება. ერთის მხრივ, ენამ, როგორც ფენომენმა, განიცადა ლინგვისტური ერთეულების (*linguistic unit*) კომბინატორული გაფართოება, ხოლო მეორეს მხრივ, უწყვეტად მიმდინარეობს ენობრივი რაციონალიზმის ანუ დახვეწის პროცესი, რაც მინიმალური სინტაქსისის პირობებში მაქსიმალური სემანტიკის ჩატევას ნიშნავს.

ინტერაქტულ პროცესში ფონემათა სტრიქონული მასივის სახით რეალიზებული მეტყველების მოდელი ერთგვარი კოლური გასაღების ფუნქციის მატარებელია, რომელიც გარე მოდელის შინაგანთან ანუ ტვინის ნეირონულ ანსამბლებზე რეალიზებულ სემანტიკურ ჰიპერგრაფებთან ასოციაციაში მდგომარეობს.

როგორც ზემოთ იყო აღნიშნული, აზროვნებისა და მეტყველების მოდელებს შორის შეუსაბამობა მათი გეომეტრიების სირთულეთა ხარისხების მკვეთრ სხვაობაშია. სტრიქონული სტრუქტურის მორფემაში მნიშვნელობა ენიჭება არა მარტო ფონემათა შემადგენლობას, არამედ მათ სტრუქტურულ მოწესრიგებულობასაც. ფონემათა ერთი და იგივე შემადგენლობა განსხვავებული თანმიმდევრობის შემთხვევაში (მაგალითად, „თმა“ და „მთა“) აბსოლუტურად სხვადასხვა სემანტიკურ ჰიპერგრაფებთან ასოცირდება. გარდა ამისა, მარტივი, ერთ სიტყვიანი წინადადება-ზმნა: „წვიმს“ რთული ტოპოლოგიის სემანტიკურ ჰიპერგრაფს ქმნის.

ერთი შეხედვით სინტაქსურად თითქოს უმნიშვნელო სასვენი ნიშნები კარდინალურად ცვლიან ტექსტის შინაარსობრივ დატვირთვას. ასე, მაგალითად: „Казнить, нельзя миловать“ და „Казнить нельзя, миловать“. ანალოგიური ვეექტი გააჩნია ზეპირმეტყველების დროს ინტონაციასაც. განხილულ მაგალითში მოყვანილ იმპერატივს ინფინიტივის ფორმა გააჩნია, რაც თარგმნის დროს შეცდომის საფუძველი შეიძლება გახდეს.

გარდა ამისა, სინონიმებისა და ომონიმების არსებობა სხვადასხვა სიტყვათშეხამების შემთხვევაში მკვეთრად ზრდის სემანტიკური ფრაგმენტების სიმრავლეს. პრობლემას ისიც ამძიმებს, რომ სხვადასხვა ენას მკვეთრად გასხვავებული სინტაქსური წყობა და გრამატიკული წესები გააჩნიათ, რაც სემანტიკასთან მიმართებაში მრავალ კაზუსს ქმნის.

### **3. ტექნოლოგიური ასპექტები**

3.1. ინტერნეტის ყველა სტანდარტული ბრაუზერი იყენებს HTML –ენაზე დაფუძნებულ დოკუმენტების წარმოდგენის ხერხებს, მათ შორის, ტექსტების სემანტიკურ ქსელებში კონვერტირების თვალსაზრისით. შესაბამისი პროგრამული სისტემების შექმნისათვის, ობიექტის სემანტიკური აღწერის აუცილებლობამ განაპირობა ობიექტ-ორიენტებული მიდგომის შექმნა, თუმცა აღნიშნული ტექნოლოგია მაინც ვერ უზრუნველყოფს მოდელის მოქნილობას, კერძოდ, თვითგანსწავლის პროცესში ცოდნის მოდელის ადექვატურ გადაწყობას. ეს უკანასკნელი ხორციელდება მხოლოდ ფრეიმების დონეზე ცოდნის ბაზაში სემანტიკური ტრანზაკციების მეშვეობით. ფრეიმებს შორის არსებული სუსტი ეპისტემიური და მით უმეტეს ვერტიკალური ასოციაციურობა გამორიცხავს ცოდნის მოდელის ჰოლიზმს.

3.2. საგნობრივი არეს ლოგიკური მოდელის დაპროექტების სტადიაზე ეფექტურად გამოიყენება UML ტექნოლოგია, რაც გარკვეულ იმიტაციას ქმნის სემანტიკურ ასპექტშიც, მაგრამ აზროვნებისა და მეტყველების ზემოსხეხებული გეომეტრიების პრინციპული შეუსაბამობა ჯერჯერობით დაუძლეველ ბარიერად რჩება კონვერტორული პროდუქტის შექმნის გზაზე.

3.3. თუ ჯერ კიდევ ახლო წარსულში მონაცემთა ბაზების ტექნოლოგიების სფეროში დომინირებდა რელაციური ტექნოლოგია, ამ უკანასკნელ პერიოდში სულ უფრო სწრაფად ვითარდება XML ტექნოლოგია, რომელსაც რიგი უპირატესობები გააჩნია მის წინამორბედთან მიმართებაში, კერძოდ ნებისმიერი დოკუმენტის აღწერის მოხერხებულობისა და დინამიურობის თვალსაზრისით, რასაც ვერ ახერხებს რელაციური ენა. მონაცემთა წარმოდგენა დოკუმენტის სახით ადამიანისათვის არის ბუნებრივი წარმოდგენა, ხოლო მისი წარმოდგენა ცხრილის სახით კი – ხელოვნურია. რელაციური ცხრილი, საუკეთესო შემთხვევაში, შეიძლება იყოს მხოლოდ დოკუმენტის ფრაგმენტი. გარდა ამისა, რელაციური ცხრილები შესაძლებელია უშუალოდ აისახოს XML-ში, მაგრამ არა პირიქით. რელაციური ბაზების ზემოაღნიშნულმა ნაკლოვანებებმა თავის დროზე გამოიწვია ობიექტ-ორიენტებული ბაზების ტექნოლოგიის შექმნის აუცილებლობა, რომელიც, სამწუხაროდ, ჯერ კიდევ ვერ ჩაითვლება დასრულებულად.

#### 4. ახალი ონტოლოგიური მიდგომა

ონტოლოგიური მიდგომა წარმოადგენს ბუნებრივ ენაზე ტექსტების ინტერპრეტაციის საშუალებას. ფილოსოფიიდან ნასესხები ონტოლოგიის ცნება ფართოდ გამოიყენება ხელოვნურ ინტელექტსა და ინფორმატიკაში. ხელოვნურ ინტელექტში იგი განიხილება ისეთი ცნებების კონტექსტში, როგორცაა ცოდნის კონცეპტუალიზაცია, ცოდნის წარმოდგენა და სხვ. თუმცა, მიუხედავად იმისა, რომ სულ უფრო ფართოვდება ონტოლოგიების გამოყენების არეალი, ხელოვნური ინტელექტის სფეროსთან მიმართებაში ჯერაც არაა ამ ცნების ზუსტი დეფინიცია.

ონტოლოგიის, როგორც ცოდნის კონცეპტუალიზაციის ზუსტი სპეციფიკაციის თვალსაზრისით ყოველი ცოდნის ბაზა, ცოდნაზე დაფუძნებული სისტემა ან ცოდნის აგენტი ფიქსირდება ცხადი ან არაცხადი სახის გარკვეული კონცეპტუალიზაციის მეშვეობით. ონტოლოგიების შედგენა გულისხმობს საგნობრივი სფეროს ობიექტების და მათ შორის სხვადასხვა კავშირების (ლოგიკური, მიზეზ-შედეგობრივი და ა.შ.) აღწერას.

ონტოლოგიური მიდგომის საფუძველს შეადგენს ბუნებრივ ენებზე ტექსტების ინტერპრეტაციის შესაძლებლობა, როგორც ემერჯენტული თვისებების მქონე სპეციალური ტერმინების სიმრავლე, რომელიც სისტემის კომპონენტებს შორის ურთიერთქმედების პრობლემის გადაწყვეტის საშუალებას იძლევა. ფორმალურად ონტოლოგიის მოდელი სამი ქვესიმრავლისაგან შემდგარი სიმრავლეა [7]:

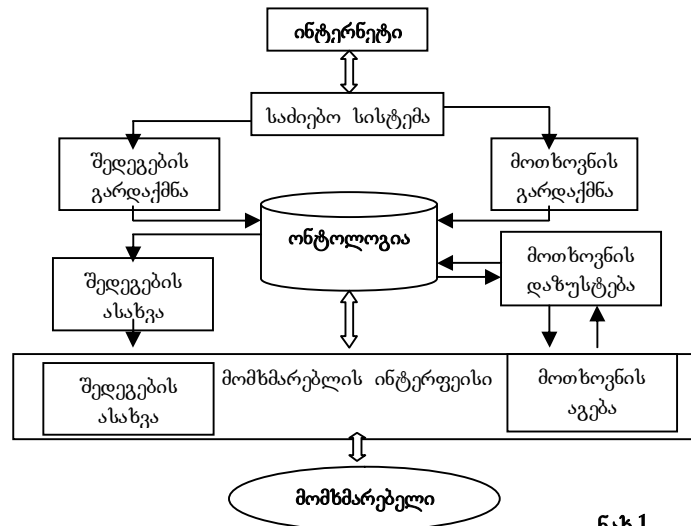
$$O = \langle T, R, F \rangle, \quad \text{სადაც} \quad (1)$$

- **T** არის საგნობრივი არეს კონცეპტების სიმრავლე;
- **R** - საგნობრივი არეს კონცეპტებს შორის დამოკიდებულებათა სასრული სიმრავლე;
- **F** - კონცეპტებზე მოცემული ინტერპრეტაციების ფუნქციათა სასრული სიმრავლე.

კონკრეტულ ამოცანაზე თუ საგნობრივ არეზე დამოკიდებულების ხარისხის, ან კიდევ აქსიომატიზაციის დეტალიზაციის დონის მიხედვით განირჩევა შემდეგი ონტოლოგიები:

- მაღალი დონის ონტოლოგიები, რომლებიც აღწერს ყველაზე ზოგად კონცეპტებს (სივრცე, დრო, მატერია, ობიექტი და სხვ.), დამოუკიდებლად პრობლემისა;
- საგნობრივი არეს ონტოლოგიები, რომლებიც შეიცავს კონკრეტული საგნობრივი არეს ცნებებს და მათი ინტერპრეტაციების სიმრავლეს;
- ამოცანათა ონტოლოგიები, რომლებიც შეიცავს ამოცანათა შესაბამის ტერმინთა სიმრავლეს;
- გამოყენებითი ონტოლოგიები – აღწერს კონცეპტებს, რომლებიც დამოკიდებულია კონკრეტულ საგნობრივ არესა თუ ამოცანაზე.

ცნობილია, რომ Internet-რესურსების სემანტიკა მრავალფეროვანია. შესაბამისად, საჭირო ინფორმაციის ძებნა მით უფრო ადვილია, რაც უფრო სპეციფიურია საგნობრივი სფერო. ამის შედეგად პრაქტიკაში შეიძლება ერთი კონკრეტული სფეროს ონტოლოგიით შემოსაზღვრა. 1-ელ ნახაზზე მოყვანილია ონტოლოგიის საფუძველზე ძებნის ორგანიზაციის სქემა.



ნახ.1.

იმისათვის, რომ მომხმარებელმა თავის მოთხოვნაზე რელევანტური აზრის პასუხი მიიღოს, ინფორმაციის ძებნა უნდა იყოს სემანტურად ორიენტებული, რისთვისაც დიდ მასივებში ინფორმაცია ორგანიზებული უნდა იყოს სემანტიკის შემცველი ონტოლოგიების საფუძველზე. აღნიშნულ ასპექტში ტექსტის შეცნობის ობიექტს წარმოადგენს სემანტიკური ჰიპერგრაფის შესაბამისი გეომეტრიული კონსტრუქცია – აზრობრივი სახე, რომელზეც რელევანტური ფრაგმენტები მკვეთრად არის გამოხატული. სწორედ ამ ფრაგმენტების კონფიგურაციებისა და ურთიერთგანლაგების მიხედვით ხდება აზრობრივი სახის იდენტიფიცირება.

ცოდნის ბაზაში ინფორმაცია ფრეიმული ჩანაწერების ნაცვლად აზრობრივი სახეების მეშვეობით იქნება წარმოდგენილი, ხოლო პარალელური ძებნის პროცედურა სწორედ სემანტიკური ჰიპერგრაფების იზომორფულობის დადგენაში მდგომარეობს. ბაზაში ცოდნის წარმოდგენისადმი ახალი ანუ გეომეტრიული მიდგომიდან, აგრეთვე Ad hoc query-ის ენტროპიულობიდან გამომდინარე, იკვეთება საგნობრივი სფეროს ონტოლოგიების ადაპტურობის აუცილებლობა, რაც სინერგეტიკული ნეირონული ჰიპერგრაფების რეკონფიგურაციით მიიღწევა [1]. ამგვარად, ახალი თაობის ინტელექტუალური ლინგვისტური პროცესორების შექმნის ტექნოლოგიური სირთულეების თვალსაზრისით დღის წესრიგში თანდათან დგება ახალი პარადიგმის აუცილებლობა.

### 5. დასკვნა: გზა ახალი პარადიგმისაკენ

მაშინ, როდესაც თანამედროვე ლინგვისტური პროცესორები ტექსტის დამუშავებას მხოლოდ წინადადებების დონეზე ახორციელებს, შესაძლოა, ერთი შეხედვით, მართლაც ცოტა შორეულ პერსპექტივად ჩანდეს, სუპერგეომეტრიული სახეების მეშვეობით ცოდნის მოდელის რეალიზაცია, მაგრამ ნანოტექნოლოგიების მზარდი განვითარება ქმნის იმის პირობებს, რომ უახლოეს ათწლეულში უკვე გამოჩნდება მოქნილი სტრუქტურის მქონე კრისტალებზე, პოლიმერულ და ბიომოლეკულებზე რეალიზებული ნეირონული მულტიპროცესორიანი ქვანტური სუპერმანქანები [7]. ეს, თავის მხრივ, აზროვნების მოდელირების, ნებისმიერი სახის ტექსტური ინფორმაციის ერთიანი სემანტიკური სახით წარმოდგენის წინაპირობა იქნება, რაც თვისობრივად

გარდაქმნის ადამიანის ინტელექტუალური საქმიანობის სფეროს, კერძოდ გადაწყვეტილების მიღების პროცესის მონიტორინგის ასპექტში. ამგვარად, ახალი თაობის ინტელექტუალური ლინგვისტური პროცესორების შექმნა მჭიდროდ იქნება დაკავშირებული ხელოვნური ინტელექტის (კერძოდ, სახეთა შეცნობის მეთოდების, მათ შორის პოპულაციური კოდირების), გრაფების თეორიის (განსაკუთრებით, ჰიპერგრაფების იზომორფულობის), ინფორმაციის თეორიის (კერძოდ, არამკაფიო ენტროპიის), აგრეთვე ინფორმაციისა და თვითორგანიზაციის ფენომენის შესახებ შედარებით ახალი მეცნიერების – ინფორმოდინამიკის, ხელოვნური ნეირონული ქსელების, ოპტიმიზაციის ევოლუციური მეთოდების შესაბამის მეცნიერებათა შემდგომ განვითარებასთან.

### **ლიტერატურა**

1. Вяткин В.Б., Синегетическая теория информации, Наука и оборонный комплекс, Мат.Межрег. научно-практической конференции. Екатеринбург: УрО РАН, 2002.
2. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001.
3. Гогичаишвили Г.Г. Автоматизация принятия решений в системах управления. „Мецნიერება“, Тбилиси, 1985
4. Шалфеева Е.А. Классификация структурных свойств онтологий. Искусственный интеллект. 2005. № 3. 67-77 с.
5. Шамис А.Л. Принципы “интеллектуализации” машинного распознавания изображений и их реализация в системах оптического чтения текстов ABBYY FineReader и FormReader, Новости Искусственного интеллекта, № 1, 2002
6. გოგინაიშვილი გ., ჩხაიძე ლ., ვეკუა მ. მარტივი წინადადებების სინტაქსური ანალიზატორი მართვის ავტომატიზებულ სისტემებთან დიალოგისათვის შეზღუდულ ქართულ ენაზე. სტუ შრ.კრ. „მართვის ავტომატიზებული სისტემები“. №4 (437), თბილისი, 2001 წ.
7. Eric K. Drexler Engines of Creation. Coming Era of Nanotechnology, ISBN 0-385-19973-2.

### **LINGUISTIC PROCESSORS: THE SURVEY ANALYSIS**

Meparishvili Badri<sup>1</sup>, Meparishvili Tamar<sup>2</sup>, Janelidze Gulnara<sup>1</sup>

1. Georgian Technical University,
2. Tbilisi State University, Georgia

### **Summary**

This paper is concentrated to the new ways of development of the linguistic processors, and to the comparative analysis of the modern knowledge representation based on the ontology, HTML technology and XML technology. The description of the fonctionnement of the human brain from the viewpoint of the shape recognition, of the treatment of texts. The semantic hypergraph as a new approach of the solution of represented problems is proposed. In conclusion the realisation of the considered semantic models on the quantum supercomputer is proposed as the draft of the future paradigm of the development intellectual systems.

### **ЛИНГВИСТИЧЕСКИЕ ПРОЦЕССОРЫ: ОБЗОРНЫЙ АНАЛИЗ**

Мепаришвили Б.<sup>1</sup>, Мепаришвили Т.<sup>2</sup>, Джanelidze Г.<sup>1</sup>

- 1-Грузинский Технический Университет
- 2-Тбилисский Государственный Университет

### **Резюме**

Преставленная статья посвящена проблемам развития лингвистических процессоров. Проведен сравнительный анализ возможностей современных HTML и XML технологии. Предложен новый подход решения рассмотренных проблем. Реалистическая модель текстового процессора предполагает также решение проблемы понимания естественно-язычного текста компьютером, а значит наличие как объекта понимания, так и соответствующей процедуры его распознавания. В предлагаемом подходе объектом понимания является искусственная конструкция – понятийный образ. В заключении реализация предложенных идей дана в рамках будущей парадигмы развития интеллектуальных систем.