

# Technological peculiarity of knowledge extraction for logical-analytical systems

Igor P. Kuznetsov, Elena B. Kozerenko, Mikhail M. Charnine

Institute for Informatics Problems of the Russian Academy of Sciences, Moscow, Russia

**Abstract** *The paper is dedicated to constructing of the new classes of expert and logical-analytical systems, based on the knowledge structures. For this the means of knowledge representation (the extended semantic networks - ESN) and the tools of their processing (the language of logical programming DEKL) have been designed. They have been used as basis for creating the new technologies, which provide the following functions: the automatic extraction of the knowledge from natural language texts, forming the Knowledge Base and the solution of the most complex problems of the logical-analytical processing by transformation and comparison of the knowledge structures. On this basis many intellectual systems for different applications have been designed.*

**Keywords** semantics, natural language, linguistic processor, knowledge extraction, named entities

## 1 Introduction

The present significance of named entities extraction pertains to the growing amount of information available in unstructured form [1-20]. The existing Internet largely consists of unstructured documents. Knowledge contained within these documents can be made more accessible for machine processing by means of transformation into relational form. An intelligent processing is proposed to transform unstructured data into something that can be reasoned with [2,5,11,13]. A typical application of entities extraction is to scan a set of documents written in a natural language and populate a database with the information extracted. More prospective approach consists of using Knowledge Base. It proposes the development of new technology including the extraction of knowledge structures and organization of their processing in Knowledge Base [3,4,6].

The distinctive features of our technology are as follows:

1. Extraction from the texts of knowledge structures (not only separate named entities) that represent the links of named entities and their participation in actions and events.

2. For the knowledge extraction the unique semantic-oriented language processors (LP) are designed. Processor LP provides the deep analysis of NL-texts and revealing set of objects together with their structures.

3. Processor LP is controlled by the linguistic knowledge, which are declarative structures (on extended semantic networks - ESN) and which provide the quick tuning of LP to subject area and language - Russian and English.

4. Linguistic knowledge consists of the rules, which provide the high degree of selectivity in the entities extraction and elimination of collisions during their

application. Rules provide the minimization of noise and losses, that is the high degree of completeness and accuracy.

5. The knowledge structures and means of their processing (intellectual language DEKL) were designed as the united tools, oriented at the tasks of linguistic analysis, semantic search, logical-analytical processing and the expert solutions. Using this tools considerably facilitates the development of applied intellectual systems.

Technology of knowledge structure extraction and processing have been used for construction of new classes of analytical systems [3,7,12]: "Criminal", "Analytic", "AntiTerror", "Resume" etc. [<http://IpiranLogos.com/en/Systems/>].

## 2 Knowledge Structures

### 2.1 Named Entities

In our technologies Named Entities (NE) are extracted from the documents on Natural Language (NL) and presented in the Knowledge Base (KB) as the fragments of the extended semantic network (ESN). The arguments of fragments are the collections of normalized words, numbers and signs, which reflect essence NE and indicate to its type. For example, the fragment

FIO(IVANOV,IVAN,IVANOVICH,1957/1+)

represents the person *Ivanov Ivan Ivanovich* 1957 year of birth. The entity type is indicated by the constant FIO. Every fragment has a unique code (sign "1+"), which corresponds to all information of the fragment and which may be on the argument places of other fragments (sign 1-). It is the main difference of the concept "fragment" from the classical concept "predicate". The network ESN consists of the set of fragments. Their order is arbitrary.

In our systems more than 40 types of NE are extracted from NL-texts. Their quantity depends on the subject area and the tasks of users. Let us note that in KB some NE can be constitutional components of others. Connections between NE may be complicated [1,6,14].

### 2.2 Type of entities and links for extraction

The set of the entities to be extracted depends on the tasks of a user. At the same time the quality of a linguistic processor is determined by the possibilities for knowledge extraction. The linguistic processors "Criminal", "Analytic" and "Semantix" support more than 40 types of semantic entities which can be extracted automatically. Some examples of basic entities types and connections extracted by the these processors are given below:

- persons (by family name, given name and patronymic - FNP) with their role features (criminal, victim);
- the verbal description of the persons, their distinctive signs;
- address, posting information attributes;
- date(s) mentioned;
- weapon with its special features;
- telephone numbers, faxes, e-mails with their subsequent standardization;
- the means of transport with the indication of the vehicle type, its state number, color and other attributes;
- passport data and other documents with their attributes;
- explosives and narcotic substances;
- organizations, positions;
- quantitative characteristics (how many persons or other objects participated in an event);
- the numbers of accounts, sums of money with the indication of the currency type;
- terrorist groups and organizations;
- participants of terrorist groups with the indication of their roles (leader, head of, etc.);
- the armed forces, assigned for antiterrorist combat (Military Force);
- event (criminal, terrorist, biographical, and so on) with the indication of the information objects participation in them;
- time and the place of events.

Standard entities (names, dates, addresses, types of weapons and others) are reduced to one standard form. The identification of entities is performed taking into account brief designations (for example, separate surnames, patronymics, Initials), anaphoric references (indicative and personal pronouns, for example, this

person, it...) definitions and explanations (for example, the *mayor of Moscow Sabyanin* is identified with the subsequent words *mayor, Sabyanin*). An important task is the identification of entities in the entire text, the use for these purposes of indicative pronouns, brief names, anaphoric references.

## 2.3 Connections between the entities and participation in actions

*Connections and relations* between NE, extracted from the NL-texts, can be very diverse. They depend on entity types. For example, one person can be connected with another by relative and friendly relations, and also by the place of living, area of interests and so on. Actions frequently are connected with the time and the place. There can be reason-consequence and other connections between actions. In such a way the complex structures are created. For their formalization special tools of knowledge representation have been designed.

*Actions* usually are expressed in NL-texts by the tensed verb forms, nonfinite verb forms, e.g. verbal nouns, participial and adverbial constructions, gerunds. The actions are also NE, components of which can be another NE. For example, there can be those, who participate in action, or entities, on which the action is directed. Moreover, some actions may be components of others. For many applications the actions are also the significant information which requires formalization. Because the system is oriented at the deep analysis of text constructions, it extracts all actions and events with NE.

Example of action extracting by the System "Semantix" is shown on fig.1 [<http://IpiranLogos/en/Semantix/>].

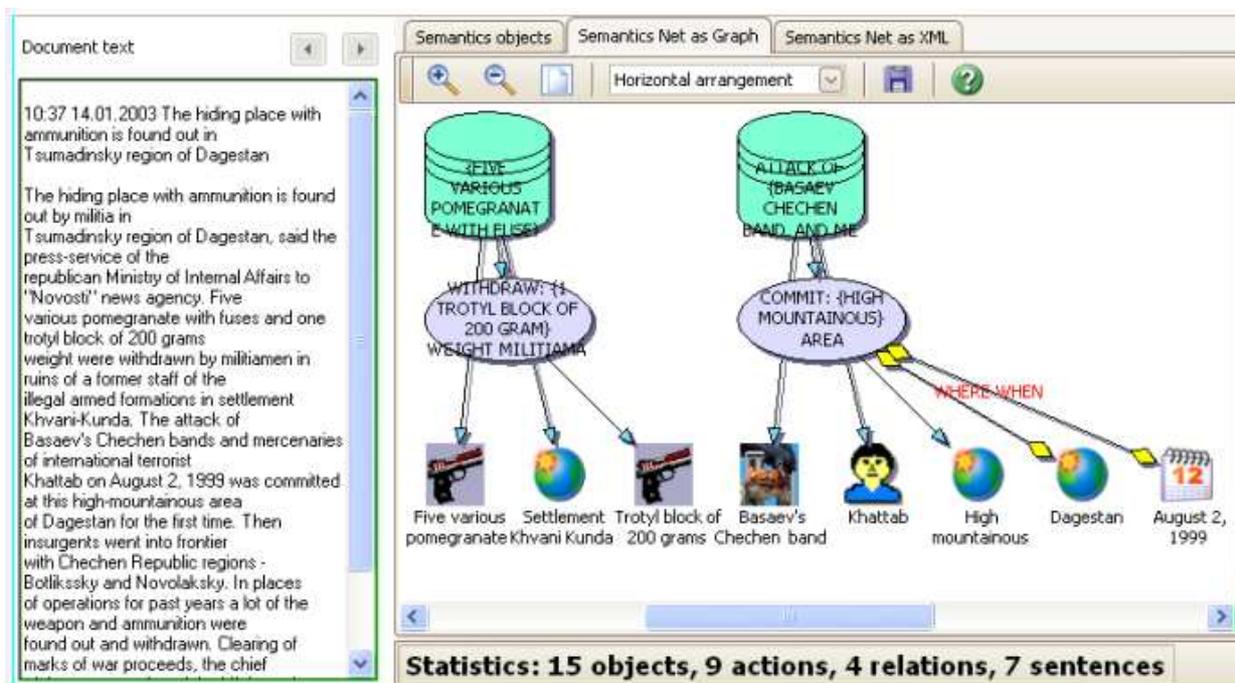


Fig.1 Example of extracted actions with entities

## 2.4 Meaningful portrait of a document

It is the formal representation of entities (NE), their properties and the connections, extracted from the text of the document. Such portraits are the structures of knowledge. As means of formalization in our technologies we use the extended semantic networks (ESN). Formalization is achieved automatically by the semantics-oriented linguistic processor, which analyzes the texts of NL-documents and transforms them into knowledge structures [3,10].

A set of meaningful portraits (together with index files) comprise the Knowledge Base (KB) where various types are provided of semantic search and logical-analytical functions by comparison and transformation of

knowledge structures. We design the technology which provides the processing in the KB distributed within the net of computers.

The Example of text (with number 22 from file e1-02-98.TXT) :

*12:16 27.12.2002 In the Chechen Republic one of leaders of bands the Arabian mercenary Abu-Tarik is destroyed. As have informed the Ministry of Foreign Affairs of the Chechen Republic, Chechen special militia destroy the insurgent in settlement Starye Atagi of Groznensky region. In one of the houses there were found the hiding place with three sub-machine guns.*

*On some data, Abu-Tarik was involved in murder of Salikhov's family in Starye Atagi in this year.*

### Meaningful portrait of the text:

DOC\_( 22, "1-02-98.TXT", "SUMMARY; " /0+) 0-(ENG)  
DATE\_(DEC.,~27,12,HOUR,16,MINUTE/1+)  
CRIM\_GROUP(1,LEADER,OF,BAND,ARABIAN,MERCENARY/2+)  
FIO("ABU - TARIK", " ", " ", " " /3+)  
DESTROY(2-,3-/4+) 4-(22,ACT\_)  
PLACE\_(CHECHEN,REPUBLIC/5+)  
WHERE(4-,5-)  
ORGANIZATION\_(MINISTRY,OF,FOREIGN,AFFAIRS,OF,CHECHEN,REPUBLIC/6+)  
INFORM(6-/7+) 7-(22,ACT\_)  
FORCE\_(SPECIAL,MILITIA/8+)  
DESTROY(CHECHEN,8-,INSURGENT/9+) 9-(22,ACT\_)  
PLACE\_(SETTLEMENT,STARYE,ATAGI,OF,GROZNENSKY,REGION/10+)  
WHERE(9-,10-)  
WEAPON\_("SUB ",MACHINE,GUN/11+)  
FIND(1,HOUSE,HIDE,PLACE,3,11-/12+) 12-(22,ACT\_)  
PLACE\_(STARYE,ATAGI/13+)  
INVOLVE(3-,MURDER,SALIKHOV,FAMILY,13-,YEAR/14+) 14-(22,ACT\_)  
  
SENTENCE\_(22,1-/15+) 15-(1,1,19)  
SENTENCE\_(22,4-/16+) 16-(1,20,114)  
SENTENCE\_(22,7-,9-/17+) 17-(2,115,288)  
SENTENCE\_(22,12-/18+) 18-(5,289,376)  
SENTENCE\_(22,ON,SOME,DATA,14-/19+) 19-(6,377,476)

A meaningful portrait consists of the elementary fragments, arguments of which are words in the normal form (it is necessary for the search and processing). Each elementary fragment has its unique code, which is written in the form of the number with the sign + and is separated by a slash line. For example, in the fragment FIO("ABU - TARIK", " ", " ", " " /3+) the sign "3+" is its code (but "3-" is the reference to it). Fragments DOK\_(22, "1-02-98.TXT", "SUMMARY; " /0+) 0-(ENG) indicate that the meaningful portrait is built on the basis of the English-language text of document with number 22 of the file of 1-02-98.TXT", which was processed as the summary of the incidents (linguistic knowledge depends on this). The following fragments present date DATE\_(.../1+),

criminal group CRIM\_GROUP(.../2+), person's surname (name and patronymic) FIO(... /3+) and so forth. The signs "0+", "0-" and "1+", "1-" and "2+", "2-" and "3+", "3-", ... are the codes of the fragments, with the aid of which their connections and relations are assigned. Actions are represented in the form of fragments of the type DESTROY(2-,3-/4+) 4-(22,ACT\_), where it is represented as "criminal group (CRIM\_GROUP with code "2+") and person (FIO with code "3+"), are destroyed". With the aid of it is the fragment 4-(22,ACT\_) indicates that the first fragment is DESTROY(.../4+) presents the action and relates to the document with the number 22. Fragments PLACE\_(CHECHEN,REPUBLIC/5+) WHERE(4-,5-)

indicate the place of this action (WHERE). Fragments ORGANIZATION\_(.../6+) INFORM(6-/7+) 7-(22,ACT\_) represent that “organization ... was informed”.

The fragments PREDL\_(...), which correspond to the sentences play the special role. They are filled up with the words, which did not enter into the named entities (in this example they are absent), or with the codes of entities themselves. To these fragments the indicators of their position in the text are added. For example, the fragment SENTENCE\_(22,7-,9-/17+) 17-(2,115,288) represents the fact that the objects with codes “7-” (corresponding to the action “inform”), “9-” (corresponding to the action “destroy” are located in the sentence, which begins from the 2nd line of the text of the document and they occupy the place from the 115-th to the 228-th byte. These means of positioning are necessary for the work of the reverse linguistic processor.

A set of meaningful portraits of documents are organized in the Knowledge Base. Logical reference is provided with the aid of the rules IF... THEN (productions) of the language DECL, which are the basis for decision of logical-analytical tasks.

### 3 Semantic-oriented linguistic processor

Semantics-oriented linguistic processor comprises the following components.

#### 3.1 The component of *lexical and morphological analysis (LMA)*

It extracts words and sentences from the text, performs lemmatization of words (normal form establishment) and constructs the semantic network presenting the space structure of text (SpST), which reflects the sequence of words, their basic features, beginnings of sentences and the presence of space character lines. The component LMA uses a two-level general ontology and a special collection of subject dictionaries (the dictionary of countries, regions of Russia, names, forms of weapons, and other items specific for the supported domains). The component performs semantic grouping of the words and assigns them additional semantic attributes [9].

#### 3.2 The component of *syntactic-semantic analysis (SSA)*

It converts one semantic network (SN) into another which represents the semantic structure of text (SemST) the, i.e., the relevant semantic entities and their connections [6,12]. The SemST is called the meaningful portrait of document. It comprises knowledge structures of the knowledge base which serves the basis for implementing different forms of semantic search: the search by features and connections, the search for the entities connected at different levels, the search for

similar persons and incidents, the search by distinctive characteristics (with the use of ontology).

The component SSA is controlled by the linguistic knowledge (LK), which determines the process of text analysis. LK includes the special contextual rules which ensure the high degree of selectivity with the extraction of entities and connections [<http://www.ipiranlogos.com/english/topics/topic3-e.htm>].

The functions of this component are the following:

- Extraction of entities from the flow of NL documents: persons, organizations, actions, their place and time, and many other relevant types of entities.
- The establishment of connections between entities. For example, persons are connected with organizations (PLACE\_OF\_WORK), by addresses (LIVES, REGISTERED). Or figurants of criminal events are connected with such entities as the type of weapon, drugs (TO HAVE).

- The analysis of finite and nonfinite verbal forms with the identification of the participation of entities in the appropriate actions. For example, one figurant gave the drugs to another figurant, and this is the fact linking them.

- The establishment of the connections of actions with the place and time (where and when some action or event occurred).

- The analysis of the reason-consequence and temporary connections between actions and events.

#### 3.3 *Expert system component (ES)*

On the basis of semantic networks the new knowledge pieces are constructed in the form of additional fragments (ESN). For example, the component ES extracts the field of a person's activity (in accordance with the assigned classifier) from the text of resume for each autobiography. The person's experience in his field is evaluated. The correlation of a criminal incident to the specific type is accomplished with the analysis of the criminal actions of ES: the following facts are revealed - the nature of crime, the method of its accomplishment, the instrument of crime, and so forth (in accordance with the classifiers of the criminal police).

3.4 *Reverse linguistic processor* which converts the meaningful portrait of document (semantic network) into the texts on NL.

#### 3.5 *Base of linguistic and expert knowledge (KB)*

It contains the rules of the text analysis and expert solutions in the internal representation. They determine the work of the linguistic processor. Our logical-analytical systems have several such bases, which are activated depending on subject areas and user tasks.

Fig.2 presents the example of extracted named entities (without links). The actions are the kind of entities (significant objects).

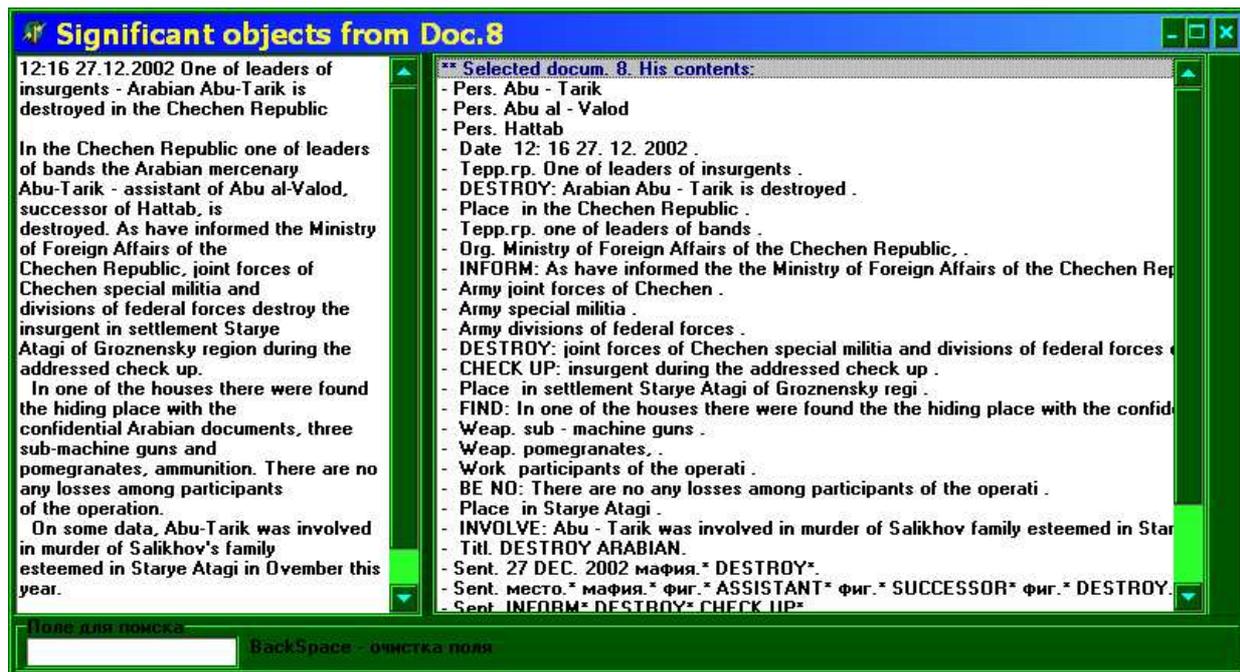


Fig.2. Example of extracted named entities

## 4 Logical analytical tasks

### 4.1 Semantic search

Semantic search is based on comparison the meaningful portrait of question and information in Knowledge Base (KB). Our technologies realize various types of semantic search

[<http://ipiranlogos.com/en/Technologies/>]:

- Search for similar entities: persons, addresses, etc.,
- Search for links (e.g. search for anonymous persons based on their word portrait),
- Search for the entities from different documents based on the indirect links.
- Answer to free questions in natural language.

### 4.2 Tasks of criminal police

System "Criminal" is based on the documents, which enter from different sources: the summary of incidents, explanatory and official notes, the notebooks of figurants (criminal persons), accusatory conclusions and other. By analysis of the documents the Knowledge Base (KB) automatically is formed [<http://ipiranlogos.com/en/Criminal/>]. Structures in KB are the basis to achieve (by methods of structural processing) the solution of the logical-analytical problems:

- the search for similar incidents and figurants according to the information in KB;

- the search for figurants by verbal portrait;
- retrieval for information on question in NL (Russian);
- the explanation of the search results;
- analysis and mapping the connections between the figurants;
- the estimation of the degree of the participation of figurants in the incident;
- the ordering figurants according to the degree of their criminal activity;
- the discovery of the organized criminal groups;
- statistical processing of information to estimate the dynamics of the criminal processes in time.

### 4.3 Tasks of recruit agency

Many services, which deal with the flows of text information, must decide the problem of their formalization: the need for representation in those forms, which are accepted in these services and within the framework which this information is used [<http://ipiranlogos.com/en/Resume/>]. For example, the important task of many recruit agencies is connected with automatic processing of autobiographical data, claim for the work of the persons (resumes, written in the arbitrary form - on NL) with the extraction of all necessary data of these persons and the forming of the computer depositories (sites, KB, tables), which provide the necessary search. The entities extracting and automatic transformation is shown on Fig. 3.

<p><b>2. IVANOVA NATALIA</b></p> <p>Telephone 935 9020(w) 248 4667(h) mobile 8 916 1266093 Moscow, Savvinskaya Nab., 4-84 10/10/1971 Objective: To obtain a position of Personal Assistant, Paralegal Assistant.</p> <p><b>Experience</b> 2002 (December)-present 000 "R.L.G." /Richemont Luxury Group Moscow, Russia Personal Assistant to Finance Director Travel arrangements Administration duties Organization of client meetings Short translations 2000 (March)-2002 (December) PricewaterhouseCoopers Moscow, Russia Partner's Assistant Preparation of Presentation in Power Point Travel arrangements Organization of client meetings Preparing materials and organizing round tables and seminars for the clients Administration duties 1997 (October)-2000 (March)</p>	<p>Last Name:-&gt; Ivanova First Name:-&gt; Natalia Middle Name:-&gt;</p> <p>Gender:-&gt; Female Birth Day:-&gt; 10 october 1971 Age:-&gt; educationId:-&gt; Grade 1 - University</p> <p>Position Name:-&gt; Obtain a position of Personal Assistant, Paralegal Assistant</p> <p>Professional Area:-&gt; 4 - Secretariat. Bec - 62.2</p> <p>Work Experience:-&gt; Grade 2 - (3-6 лет) Desireable Compensation:-&gt; Region:-&gt; Grade 1 - Moscow + region Relocation:-&gt;</p> <p>E-mail:-&gt; Home Address:-&gt; Moscow, Savvinskaya Nab. , 4 84</p> <p>Home Phone:-&gt; 2484667 Cell Phone:-&gt; Work Phone:-&gt; 9359020 Phone:-&gt; 89161266093 Url:-&gt;</p>
--	--

Fig.3. Transformation text of resume to fixed table

## 4.4 Expert systems

Expert systems on the basis of the analysis of meaningful portraits of document refer this document to the specific category (item of the classifier). In our systems two types of shells for the expert systems have been realized. The first is based on the weight coefficients of the words which correspond to the specific category. The second one is based on the presence of words in the named entities. Fig. 3 shows some examples. The "Professional area" (Secretariat – 62.2) was formed by the first type of expert system, "Region" (Moscow) was formed by second type.

## 4.5 Determination of the role functions of the entities

The technology of the determination of the new entities properties assigned implicitly is developed [16]. The

procedure of this determination based on the analysis of the knowledge structures is proposed. The task of the determination of the role functions of entities (persons, organizations and others) on the base of their descriptions is examined as the field of application. This task in general form includes all possible "estimations", "descriptions". For example, the estimation of the stability of an enterprise (according to the information from the Internet), descriptions of political figures (positive or negative depending on the statements about them in the press), the estimation of the quality of article (on the statements of users), etc. Frequently it is not said directly: this is bad, but this is good. As a rule, in the texts (NL) the events and situations are described in which one or another information object participated. On their basis the estimation is done which is often represented in the form of the new property of entities (NE). The example of determination and argumentation of role function of persons is shown on Fig. 4.

<p>12:16 27.12.2002 One of leaders of insurgents - Arabian Abu-Tarik is destroyed in the Chechen Republic</p> <p>In the Chechen Republic one of leaders of bands the Arabian mercenary Abu-Tarik - assistant of Abu al-Valod, successor of Hattab, is destroyed. As have informed the Ministry of Foreign Affairs of the Chechen Republic, joint forces of Chechen</p>	<p>Abu - Tarik - **** suspicious **** так как - Arabian Abu - Tarik is destroyed так как - Abu - Tarik was involved in murder of Salikhov family esteemed in Starye Atagi in Dvember</p>
--	--

Fig.4. Argumentation of role function of persons

## 5 Conclusion

The proposed technologies were used for construction of several intellectual analytical systems: "Criminal", "Analytic", "AntiTerror", "Monument" and others. The distinctive features of these systems are as follows: automatic extraction of knowledge structures from texts (Russian, English) and forming the Knowledge

Base which is used for realization of logical-analytical functions [13]. The ESN apparatus provides powerful representational possibilities for describing all levels of natural language, including the level of deep semantic structures, and cross-lingual correspondences.

The implemented linguistic processors were created on the basis of this approach which made it possible to manufacture design solutions for the basic problems of extracting meaningful knowledge from the texts in natural languages (Russian and English).

## References

[1] Kuznetsov I.P. Semantic Representations // Moscow: "Nauka", 1986. 290p.

[2] FASTUS: a Cascaded Finite-State Transducer for Extracting Information from Natural-Language Text. // AIC, SRI International. Menlo Park. California, 1996.

[3] Kuznetsov I.P. Methods of Processing Reports with the Extraction of Figurants and Events Features // In Dialogue'99: Proceedings of the International Workshop "Computational Linguistics and its Applications", Vol.2, Tarusa, 1999.

[4] Kozerenko, E.B. Multilingual Processors: a Unified Approach to Semantic and Syntactic Knowledge Presentation. In Proceedings of the International Conference on Artificial Intelligence IC-AI'2001 25-28, 2001. CSREA Press, 2001, pp.1277-1282.

[5] Byrd, R. and Ravin, Y. Identifying and Extracting Relations in Text // 4th International Conference on Applications of Natural Language to Information Systems (NLDB). Klagenfurt, Austria, 1999.

[6] Kuznetsov I.P., Matskevich A.G. The System for Extracting Semantic Information from Natural Language Texts // Proceedings of the Dialog International Workshop "Computational Linguistics and its Applications", Vol.2, Moscow: Nauka, 2002.

[7] Kuznetsov, I., Kozerenko, E. The system for extracting semantic information from natural language texts // Proceeding of International Conference on Machine Learning. MLMTA-03, Las Vegas US, 23-26 June 2003, p. 75-80.

[8] Kuznetsov I.P. Natural Language Texts Processing Employing the Knowledge Base Technology // Sistemy i Sredstva Informatiki, Vol.13, Moscow: Nauka, 2003, pp. 241-250.

[9] Somin N.V., Solovyova N.S., Charnine M.M The System for Morphological Analysis: the Experience of Employment and Modification // Sistemy i Sredstva Informatiki, Vol. 15 Moscow: Nauka, 2005, pp. 20-30.

[10] Kuznetsov I.P., Matskevich A.G. The English Language Version of Automatic Extraction of

Meaningful Information from Natural Language Texts // Proceedings of the Dialog-2005 International Conference "Computational Linguistics and Intelligent Technologies", Zvenigorod, 2005pp. 303-311.

[11] Cunningham, H. Automatic Information Extraction // Encyclopedia of Language and Linguistics, 2nd ed. Elsevier, 2005.

[12] Kuznetsov I.P., Matskevich A.G. Semantics Oriented Linguistic Processor for Automatic Formalization of Autobiographical Data // Proceedings of the Dialog-2006 International Conference "Computational Linguistics and Intelligent Technologies", Bekasovo, 2006, pp. 317-322.

[13] Web site "Knowledge extraction for Analytical Systems": <http://Iranlogos.com/english/>

[14] Kuznetsov I.P., Efimov D.A., Kozerenko E.B. Tools for Tuning the Semantix Processor to Application Areas // Proceedings of ICAI'09, Vol. I. WORLDCOMP'09, July 13-16, 2009, Las Vegas, Nevada, USA. - CRSEA Press, USA, 2009. P. 467-472.

[15] Kuznetsov I.P., Kozerenko E.B., Matskevich A.G. *Deep and Shallow Semantic presentations in Intelligent Fact Extractors* // Proceedings of ICAI'2010 Las Vegas, USA, June 14-17, 2010, CRSEA Press, 2010.

[16] Kuznetsov, I.P., Kozerenko E.B. *Semantic Approach to Explicit and Implicit Knowledge Extraction* // Proceedings of ICAI'11, WORLDCOMP'11, July 18-21, 2011, Las Vegas, Nevada, USA. - CRSEA Press, USA, 2011.

[17] Clark P., P. Harrison, and J. Thompson. A Knowledge-Driven Approach to Text Meaning Processing // Proceedings of the HLT-NAACL 2003 Workshop on Text Meaning, 2007. P. 1-6.

[18] Gildea D. and M. Palmer. The necessity of syntactic parsing for predicate argument recognition. In Proceedings of the 40th Annual Conference of the Association for Computational Linguistics (ACL-02), Philadelphia, PA, 2002. P. 239-246.

[19] Pasca M. and B. Van Durme. What You Seek is What You Get: Extraction of Class Attributes from Query Logs // Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07), 2007. P. 2832-2837.

[20] Punyakanok V., D. Roth, and W. tau Yih. The Importance of Syntactic Parsing and Inference in Semantic Role Labeling // Computational Linguistics 34(2), 2008. P. 257-287.