References

[Guarino, 1998] Guarino N. Formal Ontology and Information Systems. In Proceeding of International Conference on Formal Ontology in Information Systems (FOIS'98), N. Guarino (ed.), Trento, Italy, June 6-8, 1998. Amsterdam, IOS Press.

[Studer et al, 1998] Studer R., Benjamins V.R., Fensel D. Knowledge Engineering: Principles and Methods. In Data & Knolwedge Engineering, 1998, 25, p. 161-197.

[Wielinga et al, 1994] Wielinga, B., Schreiber A.T., Jansweijer W., Anjewierden A. and van Harmelen F. Framework and Formalism for Expressing Ontologies (version 1). ESPRIT Project 8145 KACTUS, Free University of Amsterdam deliverable, DO1b.1, 1994.

[van Heijst et al, 1996] van Heijst G., Schreiber A.Th., Wielinga B.J. Using Explicit Ontologies in KBS Development. In Intern. Jornnal of Human and Computer Studies, 1996, 46 (2-3), pp. 183-292.

[Artemjeva et al, 1995] Artemjeva I.L., Gavrilova T.L., Kleshchev A.S. Domain Models with Elementary Objects. In Scientific-Technical Information, Series 2, 1995, № 12. P. 8-18 (in Russian).

[Artemjeva et al, 1996] Artemjeva I.L., Gavrilova T.L., Kleshchev A.S. Logical Relationship Systems with Elementary Objects. In Scientific-Technical Information, Series 2, 1996, № 1, pp. 11-18 (in Russian).

[Artemjeva et al, 1997a] Artemjeva I.L., Gavrilova T.L., Kleshchev A.S. Logical Domain Models of the Second Order. In Scientific-Technical Information, Series 2, 1997, № 6, pp. 14-30 (in Russian).

[Artemjeva et al, 1997b] Artemjeva I.L., Gavrilova T.L., Kleshchev A.S. Logical Relationship Systems with Parameters. In Scientific-Technical Information, Series 2, 1997, № 7, pp. 19-23 (in Russian).

[Kleshchev et al, 1998] Kleshchev A.S., Artemjeva I.L., Gavrilova T.L., Surov V.V. Application of Logical Relationship Systems for Expert System Development. In Appl. of Advanced Information Technologies: Proc. of the Forth World Congress on Expert Systems, 16-20 March 1998, Mexico City Cognisant Communication Corporation, 1998, vol.1: 500-510.

Authors' Information

Alexander S. Kleshchev – kleschev@iacp.dvo.ru

Irene L. Artemjeva – artemeva@iacp.dvo.ru

Institute for Automation & Control Processes, Far Eastern Branch of the Russian Academy of Sciences

5 Radio Street, Vladivostok, Russia

INFORMATION PROCESSING IN A COGNITIVE MODEL OF NLP

Velina Slavova, Alona Soschen, Luke Immes

Abstract: A model of the cognitive process of natural language processing has been developed using the formalism of generalized nets. Following this stage-simulating model, the treatment of information inevitably includes phases, which require joint operations in two knowledge spaces – language and semantics. In order to examine and formalize the relations between the language and the semantic levels of treatment, the language is presented as an information system, conceived on the bases of human cognitive resources, semantic primitives, semantic operators and language rules and data. This approach is applied for modeling a specific grammatical rule – the secondary predication in Russian. Grammatical rules of the language space are expressed as operators in the semantic space. Examples from the linguistics domain are treated and several conclusions for the semantics of the modeled rule are made. The results of applying the information system approach to the language turn up to be consistent with the stages of treatment modeled with the generalized net.

Keywords: Cognitive model, Natural Language Processing, Generalized Net, Language Information System

ACM Classification Keywords: I.2.7 Natural Language Processing;

Introduction

Natural language processing (NLP) is a complex cognitive function, representing a complicated subject for modeling and formal description. The trial in this wok is to elaborate a reliable formal model of NLP in order to propose a tool for analyzing the process and to examine the possibilities for further implementations.

The formal model of NLP, presented here, is elaborated using a cognitive science approach. The intention is to take into consideration as much as possible the essential cognitive principles that most cognitive scientists agree with:

1). The mental system has a limited capacity - the amount of information that can be processed by the system is constrained.

2) A control mechanism is required to oversee the encoding, transformation, processing, storage, retrieval and utilization of information.

3) The constructing of meaning is a dynamic process resulting of a two-way flow of information – the flow, gathered through the senses (Bottom-up processing) and the flow of the information, which is stored and classified in the memory (Top-down processing)[1].

4) The human organism has been genetically prepared to process and organize information in specific ways. In the further description we'll consider that all these functions are performed within a system, called *"cognitive system"*.

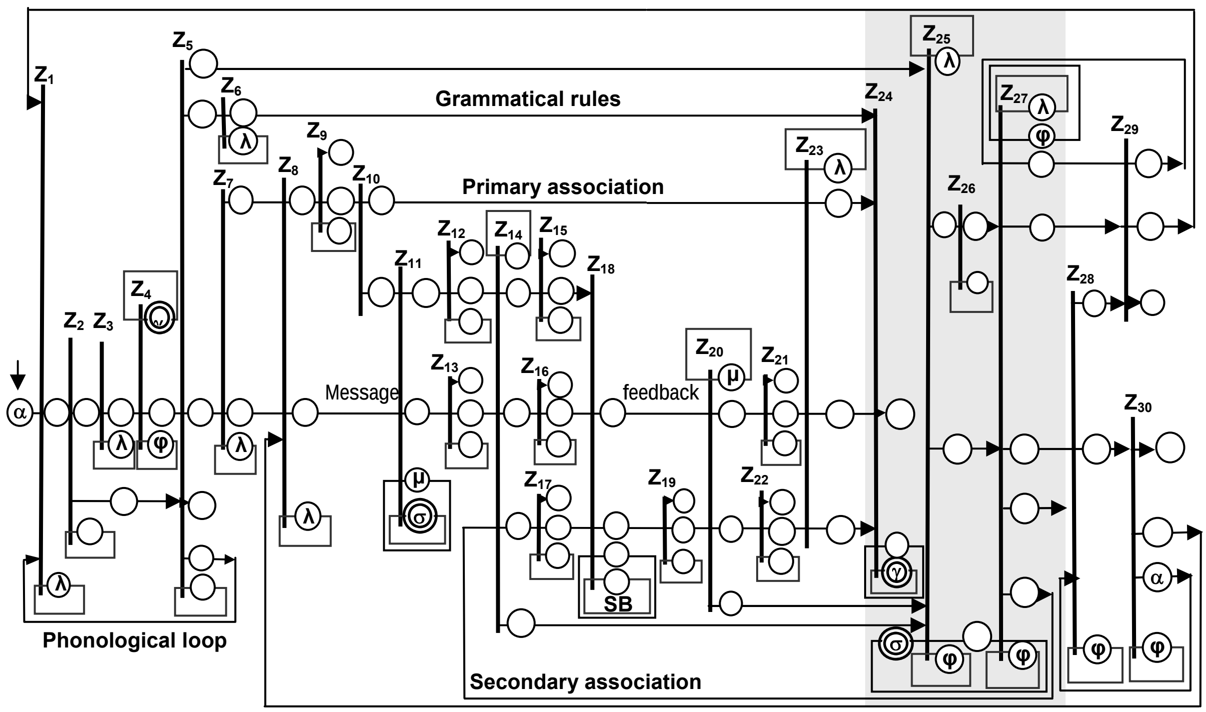

Figure 1. AGN - Generalized Net model of process of message acquisition

A formal model of NLP has been elaborated, using the mathematical formalism of Generalized Nets (GN). The obtained Net, called AGN (Figure 1) gives a formal description of the cognitive process of treatment of a language message, arriving on the input of the auditory system (Bottom), processed stage by stage and conduced

1 Concerning the cognitive aspects of language processing, the constraints on linguistic performance come mostly from the top-down information processing.

to the mind (Top) through several parallel pathways. AGN[1] corresponds to the cognitive system, providing a control mechanism for overseeing the transformation, processing, storage and retrieval of information.

AGN treats language message, consisting in sentence-fragments, formally expressed as a sequence of α-tokens. Each α-token, travelling from the input to the final position of the net, is submitted to consequent treatments, performed on the transitions Z_i. AGN transitions Z_1 – Z_{29} imitate the phases of the process of speech perception. The information, obtained by the system, is formalized as α-token's characteristics, which are acquired when crossing the transitions. The net al.ows modeling the interaction between the Bottom-up and Top-down information flows. The Top-down information flow assigns new characteristics to the "travelling up" signal. To imitate the parallel "emergence" of the gathered information of different type, the α-token splits (see for example Z_5), follows different pathways and terminates by fusing all obtained characteristics into an internal lexical and semantic representation of the message content.

The Top-knowledge, stored in Long Term Memory (LTM) is organized in two related spaces[2] – the language space (as a system of lexical units and rules) and the semantic space (the semantic representation of the world as a system of semantic primitives and rules). They have respectively two underlying structures - the word-forms graph WG, expressed by the γ-token, and the semantic net NSet, expressed by the σ-token. The result of the Top-down flow is stored in WG as "*expectation*" of word-forms[3]. Four sources of expectation are modeled. Two are related to the language - the memorized language practice, called "*primary association*" and the knowledge of *grammatical rules*. Two other sources of expectation are due to semantic activation: the listening-comprehension *message feedback* (caused directly by the word-forms in the message) and the "*secondary association*" (semantic activation, accumulated because of the sequence of the message word-forms). Tokens γ and σ are thought as structures, which elements accumulate expectation/activation. AGN has access to each of them on two places - one 'retrieval' place (Z_4 and Z_{11}) and one place for storing expectation/activation (Z_{24} and Z_{25})

The knowledge of language is represented by a number of λ-tokens, each skilled with a treatment procedure as characteristic. They are on transitions: Z_1 - with: "Segmentation procedure *Seg*"; Z_3 - with: "Phonemes recognition procedure *Rec*"; Z_6 - with: "Grammatical features and dependencies procedure *Gr*"; Z_7 - with: "Lexeme representative retrieval procedure *InL*"; Z_8 - with: "Primary association procedure *Ass1*"; Z_{23} - with: "Lexeme members retrieval procedure *Lexemize*"; Z_{25} - with: "word-forms concordance procedure *TreeBranches*"; and Z_{27} - with: "Syntax structure discovery procedure *Parse*". The cognitive processes are expressed by φ-tokens, and the "mental dictionary" (relations on WG x NSet) - by μ-tokens. All this tokens, with procedures as characteristics, are 'turning' over the corresponding transitions during the time-period of AGN functioning.

Initially, token α enters the net with characteristics "Phonological features". At transition Z_1, a λ–token "Language knowledge – prosody" skilled with the Segmentation procedure transforms the input to a "Sentence segmented into word-form segments", given to the α-token as characteristics. Transition Z_2 simulates auditory sensory memory. Transition Z_3 corresponds to the stage of phoneme recognition. Transition Z_4 corresponds to the comparison of the recognized phonetic content with the lexical knowledge retrieved from WG. Transition Z_5 accepts or rejects the retrieved word-form for further treatment (AGN is supposed to identify the input word-forms using the *expectation,* gathered in the units of WG). Transitions Z_6 to Z_{24} represent a Working Memory (WM) Sub-Net[4] where the two information flows meet[5] and generate expectation. Multiple transitions in this part are connected to LTM-tokens, producing lists of LTM knowledge (lexeme representatives, synonyms, homonyms,

[1] AGN has been presented in a series of papers in the domain of information technologies and cognitive modelling in linguistics (see for example [Kujumdjieff, Slavova, 2000]. Here we follow the numbering and the names, accepted in the complete formal description of the net, given in [Slavova, 2004].

[2] Most cognitive researchers agree on the different nature of the language knowledge and the conceptual knowledge, including their separate localization on the cortex.

[3] It is known that the capacities of speech perception do not allow capturing all the pronounced phonemes. In fact, the cognitive system constructs the missed, but 'expected' content of the message. The same top-down phenomenon is available when reading texts.

[4] This sub-net is presented in details in [Slavova, Atanassov, 2004).

[5] According to the most part of the existing in cognitive science theories and models of memory, the Top-down and the Bottom-up flows meet using Working Memory resources.

concepts, attributes etc.) The sub-net imitates limited WM resource and 'concentration' on the message content by retaining only the heads of the lists, sorted following the accumulated activation. On Z_{24} the expectation from all pathways is overlapped and stored in the elements of WG (see also figure 2, token γ on place l_{82}). Transition Z_{25} simulates the activation of NSet by the message word-forms. Transition Z_{26} is a working memory for lexical units. Transition Z_{27} expresses the process of analyzing the entire sentence. Transition Z_{28} simulates the extraction of basic semantic roles (with a feedback from the message memory content). Transition Z_{29} simulates rechecking (if the semantic roles can not be properly discovered). The α-token is finally stored with all obtained characteristics in the message memory, represented on transition Z_{30}.

The Generalized Net approach has allowed formalizing, on a high level of abstraction, the cognitive process of message acquisition. This representation allows incorporation of sub-nets and separate modules, such as databases, neuron nets etc. Such an approach starts to be used in hybrid nets in AI [see Atanassov, K., 1998].

The Problem – Parallel Language and Semantic Treatment

A big part of the treatment procedures, introduced on AGN transitions, such as *"segmentation procedure"* or *"expectation dependent retrieval from WG"*, are easy to be imagined. The problem is to conceive the procedures, which run on the transitions after the WM sub-net. The tracking of the cognitive process has lead to base them on semantic and language knowledge at once. The presented work gives further development of AGN by analyzing the procedures, which perform simultaneously in the language and the semantic space.

Transition Z_{25} (figure 2) simulates two processes, which run in parallel. The first is the activation of the semantic space by the message word-forms W and corresponds to building a mental image of W in terms of concepts and features. The second one is the detection of related words in the sentence and occurs at the moment when the grammatical features and the semantic image of W are discovered. It is supposed that the cognitive system first assembles a fractional representation of the sentence-meaning structure (coupled words for example) by consulting the semantic net for incompatibilities, as the grammatically determined word-chains have to be coherent with the meaning of the corresponding concepts and/or features. The formal description of Z_{25} is:

Z_{25} = < $\{l_{12}, l_{49}, l_{69}, l_{70}, l_{83}, l_{89}\}$, $\{l_{83}, l_{85}, l_{87}, l_{88}, l_{89}\}$,

	l_{83}	l_{85}	l_{87}	l_{88}	l_{89}
l_{12}	*false*	*true*	*true*	*false*	*false*
l_{49}	*false*	*false*	*true*	*true*	*false*
l_{69}	*false*	*false*	*false*	*true*	*false*
l_{70}	*false*	*false*	*true*	*false*	*false*
l_{83}	*true*	*false*	*true*	*false*	*false*
l_{89}	*false*	*false*	*false*	*true*	*true*

, $\vee(\wedge(l_{70}, l_{83}, l_{12}), \wedge(l_{89}, \vee(l_{49}, l_{69})))$>.

The α-token enters transition Z_{25} with the following characteristics, coming from:

position l_{12} – W, word-form assumed to be perceived (on transition Z_5) with its grammatical feathers GrFtrs;

position l_{49} - Nt ct (from the message feed back pathway) – the head of the list of semantic net elements NSet, which correspond to W (the correspondence is found on Z_{11} using the mental dictionary - μ-token);

position l_{69}, - NtSBlist - the first n of list of NSet elements, which correspond to the received up to the moment W - s, arranged following the number of their manifestations in a semantic buffer SB (secondary association).

On transition Z_{25}:

λ-token "Language knowledge - syntax and grammar" turns on place l_{83} with:

"word-forms concordance - Procedure *TreeBranches* ";

σ-token, the Semantic net NSet stays on place l_{70}:

"Semantic net elements – NSet";

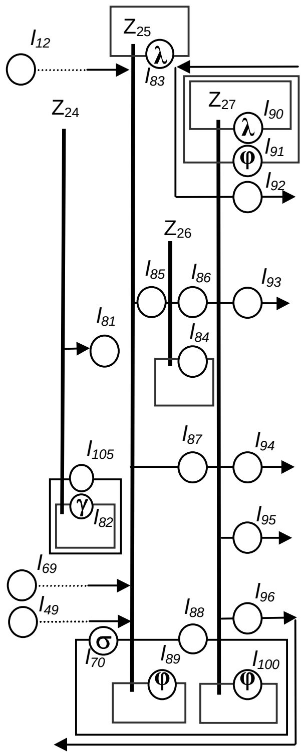

Figure 2. Two of the transitions, based on language and semantics

φ-token “Cognitive process - semantic activation” turns on place l_{89} with “Storage of activation in NSet nodes - Procedure *SemA*”;

After Z_{25} , on place l_{88} σ-token takes the characteristic:

“ANet = *SemA* (Nt ct, NtSBlist) – activation of NSet elements”

Token α obtains in place l_{87} the characteristic:

“ParSynStr = *TreeBranches* (NSet, GrFtrs) – Partial syntax structure.”

Tokens do not change their characteristics on places l_{83}, l_{85} and l_{89}.

Transition Z_{26} is WM buffer for W, queued on l_{84} and transmitted to l_{86}.

On transition Z_{27}, the following procedures are running:

λ-token “Language syntax knowledge” stays on place l_{90} with:

“Syntax structure discovery - Procedure *Parse*”;

φ-token “comparing semantics and syntax” stays on l_{91} with:

“Comparing – Procedure *Comp*”;

φ-token “focus determination” stays on place l_{100} with:

“Semantic center localization - Procedure *DetSC*”.

Transition Z_{27} expresses the mental process of analyzing the entire sentence after its last word-form has been perceived. It is assumed that two parallel processes take place at this time-moment: the sentence syntax structure is clarified and the semantic focus Nt^1 of the sentence is detected (NSet element, staying higher in NSet’s hierarchy). The brought by α-token information, acquired before entering Z_{27}, consists of: partial syntax representation, word-forms W in the lexical buffer content and activation of the corresponding nodes of NSet. It is supposed that the syntax structure of the sentence is recognized with semantic justification.

Z_{27} = < {l_{86}, l_{87}, l_{88}, l_{90}, l_{91}, l_{99}, l_{100}}, {l_{70}, l_{90}, l_{91}, l_{92}, l_{93}, l_{94}, l_{95}, l_{96}, l_{100}},

	l_{70}	l_{90}	l_{91}	l_{92}	l_{93}	l_{94}	l_{95}	l_{96}	l_{100}
l_{86}	*false*	*false*	*false*	*true*	*true*	*true*	*false*	*false*	*false*
l_{87}	*false*	*false*	*false*	*false*	*true*	*true*	*false*	*false*	*false*
l_{88}	*true*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*false*
l_{90}	*false*	*true*	*false*	*false*	*true*	*true*	*false*	*false*	*false*
l_{91}	*false*	*false*	*true*	*false*	*true*	*true*	*false*	*false*	*false*
l_{99}	*false*	*false*	*false*	*true*	*true*	*true*	*false*	*false*	*false*
l_{100}	*false*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*

$,\vee(\wedge(\vee(l_{86}, l_{99}), \wedge(l_{88}, l_{90}, l_{91})), \wedge(l_{88}, l_{100}))$ >

In places l_{93} and l_{94} the α-token obtains the characteristic:

“TRes = *Comp* (*Parse*(Buff, ParSynStr), NSet) - Obtaining complete syntax structure”

and in places l_{95} and l_{96} the α-token obtains the characteristic:

“Nt^1 = *DetSC* (NSet, ANet) – momentary semantic center”

Tokens do not change their characteristics on places l_{90}, l_{91}, l_{92}, l_{70} and l_{100}. The procedure *Comp* may have two results: 1. TCF (Tree Construction Failed); 2. TREE+ WsC (Syntax tree with W corrected - WsC).

The procedure *Comp* (*Parse*(Buff, ParSynStr), NSet) on Z_{27} is applied on the retained in a STM buffer for W and on the result of *TreeBranches* (NSet, GrFtrs), running on Z_{25}. Both procedures are based on NSet and on the grammatical features of W (discovered on Z_4).

The Language as an Information System

The generalized net model suggests that we have to formalize Z_{25} and Z_{27} through procedures, running by using in parallel semantic and language knowledge and related to the elements and to the rules in the two knowledge spaces. A sequence of questions appears concerning the simultaneous operations in the two spaces. In AGN, the correspondences between them are given by the 'mental dictionary', containing the relations between the lexical units in WG and the elements of the semantic net NSet, but the structures of the two spaces are independent. Are there rules that allow joint operations in the two spaces? Are they established on principles for mapping the structures of the two spaces? That needs to examine the human language in a general way.

Let us present human language as an Information System (IS) – a Language Information System (LIS). One of the primary goals of a human language is to assure the information exchange between individuals. Information, residing as *internal cognitive representation* of the individual H_1 is first presented as language-coded information, communicated to another individual H_2, and interpreted to *internal cognitive representation* of individual H_2. It is intriguing to provide the example of one home-made sign language, created and utilized by two deaf sisters. The used pointing gestures were found to be part of lexical terms referring to present and non-present objects, persons and places. Some gestures occupied fixed positions in sentences, apparently used as grammatical terms. Oral movements were frequently used together with manual signs, and their functions may be classified as lexical, adverbial, and grammatical (Torigoe et al., 2002). This strongly suggests the existence of the *innate* mechanism for mental representations (Hauser 2002, Chomsky 2004).

Imagine we have to construct a LIS. Let us apply the used in the technology domain procedure. For conceiving an IS, the representation "*input - treatment block – output*" is used. On its input, an IS receives *data* and *resources* and on its output - obtains *informative products*. The treatment block runs on the bases of a particular *model and method for data-processing,* which includes a number of *rules* and *operators* on data. So:

1. The *resources* of LIS are the human cognitive resources – *static* (long term memory and working memory), and *operational* (operators that human mind performs on the available operable substances).

2. Data. The mind-operable content of the data-source (figure 3) has to be transmitted as data, operable in language. The functioning of the data-source has to be presented as a model, a system of elements with determined roles, reproducing how the cognitive system operates when performing its tasks. Let us call the components on this model "semantic primitives". A structure of data-containers has to be assembled in the language and matched to the structure of the semantic primitives. Data-values have to be accorded to all distinct entities, available and operable in the source, and stored in the corresponding data-containers. This approach is well-known in the IS domain (see for example Codd, 79).

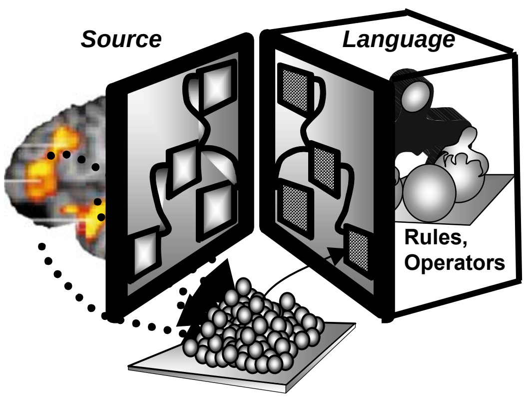

Figure 3. The language, seen as information system

3. The *treatment block* requires conceiving a set of *rules* and *operators* on data. This set has to allow generating larger units, reproducing the processing in the source. The cognitive system will operate data following these rules, so they must be expressible by means of the operators, which run on the semantic primitives.

The *final product* of the language has to create an accurate internal representation of individual H_2, who receives the LIS output. The 'decoding' is done with the active participation of H_2 (the 'Top-down" information processing), which presupposes that H_2 knows the language and possess a semantic description of the world.

The information system reasoning shows that the accurate functioning of a LIS strongly relies on the internal semantic representation. The language is constructed on the bases of the semantic primitives and the mind operators on them. The claim is that in all languages must exist interactions between the purely language features and their semantic fundamentals. This could give a basis for the joint semantic-language operations, leading to a solution of the problem concerning the procedures on Z_{25} and Z_{27}. The next step is to examine the relationships between the grammatical level and the semantic one, starting from concrete working examples.

Cognitively Based LIS on the Example of Russian

We assume that there is a common general underlying semantic scheme for all languages. Then it will follow that any grammatical rule can be represented as consisting of some semantic primitives as internal representations, which are mind-operable. We did a trial to show how LIS operates on the example of a concrete syntactic representation in one specific language. Secondary predication is a grammatical particularity in Russian, allowing variability of case marking (Instrumental/Nominative) on secondary predicates. Example:

(1) a. Maria prišla ustalaja-nom.

b. Maria prišla ustaloj-instr.

Mary arrived tired.

In Russian, we can make the notion of arriving tired, simultaneously being contrasted with some non-tired state, using the instrumental case. In the nominative case, we only know that 'she arrived tired', with no past reference to any other possible state. The semantics of Russian secondary predication has been examined a lot by the specialists in linguistics and the obtained results and explanations are not uniform.

We have followed the LIS reasoning and we constructed a database (DB) in which exists simultaneously the language level and the semantic level, with their structures, interconnected.

Examples of statements, taken from the linguistics studies of secondary predication, (53 sentences) wore stored in the table Examples. The cases in Russian are used as markers for the grammatical annotation of the examples. Data, expressing the language and semantics spaces, have been organized in tables as follows (figure 4):

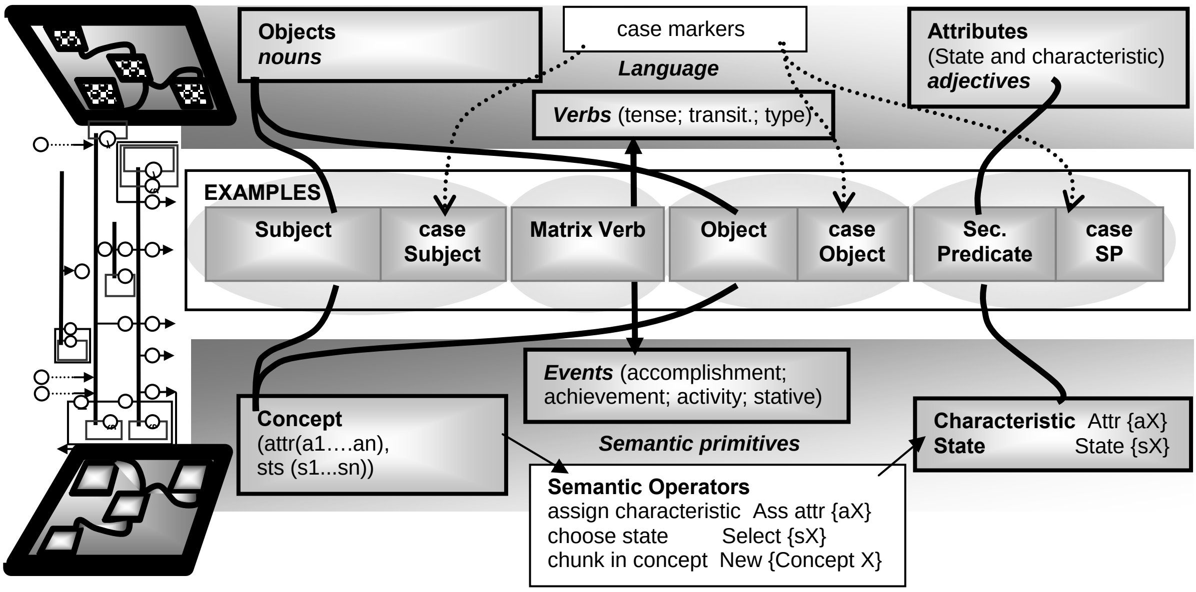

Figure 4. Language – semantics database design

Table Objects stores all *concepts* from the examples, with their names in different languages. The *attributes* - characteristics or possible states, are stored in a separate table. The verbs with their grammatical features are stored in the table 'Verbs' (they are introduced in the field "Matrix Verb" of the table Examples as foreign key). The *events*, providing the underlying semantic features of the verbs, are expressed as attributes of the verbs (foreign key). Events are stored separately with their semantic features following some of the existing classifications. With this construction of the DB, the grammatical rules of case marking can be examined independently of the events' structure. For the purposes of this analysis, we employ a revised version of the 'event structure' (Davidson 1967, Vendler1957, Verkuyl 2001), examining the language semantics primitives.

The statements are assigned two levels of representation - phases of semantic-levels-translation: The first step accords to a lexical item its basic semantic category. The categories that we used are *concept*, *characteristic*, *state* and *event.* We assume that the grammatical level, expressed by means of case markers, implies running of semantic operators. For our examples we took as basic operators the following set: "*assign characteristic*: Ass attr {aX}", "*choose state:* Select {sX}" and "*chunk in concept*: New {Concept X}".

The second representation of the statements gives the result of applying the semantic operator. Using queries over the modeled in the DB parameters, we checked several guesses about the semantic interpretation, coming from the linguistics domain. For example, as the running of queries over the events characteristics does not give any indication of changing states, so the conclusion is that the meaning of the matrix verb events is not influenced by the case marking. It is interesting to see the result of queries, which put together (taking data from the corresponding tables) all language labels (Russian and English in Table 1), the semantic markers and the verb-event information. Here are a few of the examples for transitive verbs (*Re* - the results of the first phase of semantic-levels-translation, *Se* – the semantics of the sentence, obtained after the second phase):

Table 1

	13a.			
Ex	Ja/-nom	Pokupaju	banany/-acc	spely/-instr
	I/-nom	*Buy*	*bananas/-acc*	*ripe/-instr*
Re	Ja (concept)	Pokupaju	banany (concept)	spely (state)
	I (concept)	*Buy*	*bananas (concept)*	*ripe (state)*
	attr(a1….an), sts (s1….sn)		attr(a1….an), sts (s1….sn)	
Se	Ja (concept)	Pokupaju	banany-spely (selected state)	
	I (concept)	*Buy*	*bananas-ripe (selected state)*	
	attr(a1….an), sts (s1….sn)		in state {sX}	
	Activity	(<…,<sn,sn+k,k=1>,…>)	'I buy bananas ripe.'	
	16a1			
Ex	Don/-nom	Pišet	pis'mo/-acc	ustal/-instr
	Don/-nom	*Writes*	*letter/-acc*	*tired/-instr*
Re	Don (concept)	Pišet	pis'mo (concept)	ustal (state)
	Don (concept)	*Writes*	*letter (concept)*	*tired (state)*
	attr(a1….an), sts (s1….sn)		attr(a1….an), sts (s1….sn)	state {sX}
Se	Don-ustal (selected state)	Pišet	pis'mo (concept)	
	Don-tired (selected state)	*Writes*	*letter (concept)*	
	in state {sX}		attr(a1….an), sts (s1….sn)	
	Activity	(<…,<sn,sn+k,k=1>,…>)	'Don writes letter tired.'	

It come out that the use of a canonical underlying semantic scheme of events, objects, states and attributes explains the semantics of the grammatical rule of secondary predication in a clear way:

The case marking of the secondary predicate implies meaning of a "*choose state:* Select {sX}" operator in the case of instrumental and an "*assign characteristic*: Ass attr {aX}" operator in the case of nominative. The grammatical rule of secondary predication is applied to concepts and plays a role of a "choice of state" operator without influencing the structure of the matrix verb's event. The semantic-level interpretation of all statements shows that the event structure plays its role for the meaning of the sentences in an independent way.

This clear representation of secondary predication may be implemented in several ways. We constructed a Neuron Net (figure 5), which performs the treatment of secondary predication grammatical rule, as the aim is to include further the treatment of other rules.

Our task is to model the more explicit Russian syntax. So, we choose our essential cognitive features, with just a sufficient enough neural network. The nodes of our neural network are conventional symbols: nouns, adjectives, adverbs or verbs. Activation of the node occurs when a sufficient threshold value is reached. The sum is used for AND operation, and 1 is use for OR operation. Initially, all inputs are OFF. There is no learning component.

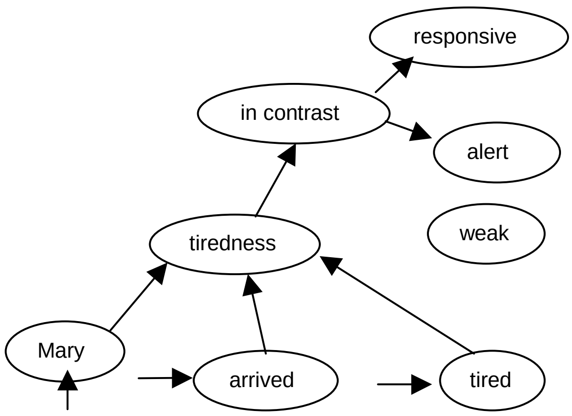

Figure 5. Neuron Net

Word order is not part of the model, but left up to the actual language. Because Russian has a more flexible word order, and the net meaning is the same for both Russian and English (as concerns ordering of words), then assume words have been entered in the right order.

The neuron net imitates: Instr. case marking on *tired* in *Mary arrived tired* triggers a state of *tiredness.* Being in a state of tiredness gives rise to an explicit state *in contrast*, then non-tired states: of: *alert*, and *responsive*.

The advantages of this representation are 1) semantics and syntax are combined; 2) knowledge engineering is much easier, including maintenance, because symbols are used, instead of dynamic numeric values. Clearly, the neuron net has to have a much richer set of semantics for a practical system.

Conclusion

The supposal that the cognitive system treats in parallel the semantic and the language knowledge space was made on the bases of two formal representations: the AGN model of the cognitive process and the representation of the language as an information system. The results of these two formal approaches are in agreement. The 'semantic' database representation of primary/secondary predication on the example of Russian was used in the analysis of the links between basic semantic units and grammar. Grammatical rules of the language space are expressible as operators in the semantic space. Some important linguistics conclusions wore made on this base.

It is interesting to analyze the content of Table 1. It became clear that for intransitive verbs the *choice of state* operator is applied always to the state of subject, but for transitive verbs it can also be applied to the state of object. Obviously in the statement 13a the state 'ripe' can not be accorded to the subject 'I' and in 16a1 the state 'tired' is not for the object 'letter'. But these two statements are absolutely correct, they are not ambiguous and in use in both languages. From the point of view of AGN treatment, on Z_{25} the procedure *TreeBranches* has to 'attach' the word-form "ripe" to "bananas" and the word-form "tired" to "I", as it consults the semantic space NSet, where the concepts and their attributes are known. The further development of this work necessitates formalizing in a detailed way the AGN-s part Z_{25} Z_{27}, associated to LIS on order to implement this part of the model.

Bibliography

[Atanassov, 1998] K. Atanassov. Generalized Nets in Artificial Intelligence. Vol. 1: Generalized nets and Expert Systems. "Prof. M. Drinov" Academic Publishing House, Sofia

[Becker, Immes, 1987] L.A. Becker, L. Immes. An approach to automating knowledge acquisition for Expert Systems: annotated traces -> Diagnostic Hierarchies. In: ACM Conference on Computer Science, p133-137.

[Chomsky, 2004] N. Chomsky. The Generative Enterprise Revisited. Mouton de Gruyter.

[Codd, 1979] E. F. Codd. Extending the Database Relational Model to Capture more Meaning. In: ACM TODS, Vol 4, n 4.

[Davidson, 1967] D. Davidson. The Logical Form of Action Sentences. In: N. Rescher, ed., The Logic of Decision and Action, Pittsburgh University Press, Pittsburgh.

[Hauser et al., 2002] M. Hauser, N. Chomsky and W.T. Fitch. The Faculty of Language: What is it, who has it, and how did it evolve? In: Science Vol. 298.

[Kujumdjieff, Slavova, 2000] T. Kujumdjieff, V. Slavova. An association-based model of natural language comprehension. In: International Journal "Information theories & applications", Vol. 7 Number 4, 169-174.

[Torigoe et al., 2002] Torigoe, Takashi, Takei, Wataru. Descriptive Analysis of Pointing and Oral Movements in a Home Sign System. In: Sign Language Studies 2.3, Spring 2002;

[Slavova, 2004] V. Slavova. A generalized net for natural language comprehension. In: Advanced Studies in Contemporary Mathematics, vol 8, Ku-Duk Press, 131-153.

[Slavova, Atanassov,2004] V. Slavova, K. Atanassov. A generalized net for working memory and language. In: Text Processing and Cognitive Technologies, VII International conference «Cognitive modeling in linguistics», Varna, 90-101;

[Soschen, 2003] A. Soschen. On Subject and Predicates in Russian. Ph.d. Dissertation, Univ. Ottawa.

[Vendler, Zeno.,1957] Vendler, Zeno. Verbs Verbs and Times. In: Philosophical Review 66, 143-160.

[Verkuyl, H.J., 2001] H.J. Verkuyl. Aspectual Composition: Surveying the Ingredients. In: Proceedings of the Utrecht Perspectives on Aspect Conference December 2001, OTS.

Author's Information

Velina Slavova – New Bulgarian University, Department of Computer Science, 21 Montevideo str., 1618 Sofia, Bulgaria, e-mail: vslavova@nbu.bg

Alona Soschen – Massachusetts Institute of Technology, MIT Department of Linguistics and Philosophy 77 Massachusetts Ave. 32-D808Cambridge, MA 02139-4307, USA, e-mail: soschen@mit.edu

Luke Immes – Multisegment Co., 20 Williamsburg Court, suite 14 Shrewsbury, MA 01545, USA, e-mail: li@multisegment.com

MATHEMATICAL MODEL OF RE-STRUCTURING COMPLEX TECHNICAL AND ECONOMIC STRUCTURES

May Kornijchuk, Inna Sovtus, Eugeny Tsaregradskyy

Abstract: Research and development of mathematical model of optimum distribution of resources (basically financial) for maintenance of the new (raised) quality (reliability) of complex system concerning, which the decision on its re-structuring is accepted, is stated. The final model gives answers (algorithm of calculation) to questions: how many elements of system to allocate on modernization, which elements, up to what level of depth modernization of each of allocated is necessary, and optimum answers are by criterion of minimization of financial charges.

Keywords: system, re-structuring, quality, reliability.

ACM Classification Keywords: I.6.3 Simulation and Modeling: Applications

Introduction

By development of new complex systems, and increase of their efficiency while in service the important factor of increase of adequacy and reliability of mathematical models an estimation of a level of their reliability is ability of the description, formalization and the account in these models of an opportunity of management of reliability [1]. The increment of reliability *u* due to rational management of reliability is achieved by perfection of algorithm of a system's mode of operation variations, a variation of actions on technical and to preventive maintenance, because that reduction of failure rate after rational procedure of procedural works depends on a level of optimization of this procedure. Reduction of intensity of a refusal's stream, change of its probable structure of limited after action can be achieved also by special modes of external influences. So, for example, separate kinds of integrated circuits at a radioactive irradiation sharply raise accuracy of a presence of parameters in necessary borders [1]. However the time of their life essentially decreases. Realization of such procedure when the system carries out the important and responsible task nevertheless can be quite justified. Value of such task allows neglecting reduction of general time of life of an element due to strict preservation of parameters in