

*Linguistics*

## The Structure of Interactive Language Model Algorithms Based on the Net System

Giorgi Chikoidze\*, Ana Chutkerashvili\*, Nino Javashvili\*

\*Archil Eliashvili Institute of Control Systems of Georgian Technical University

(Presented by Academy Member Avtandil Arabuli)

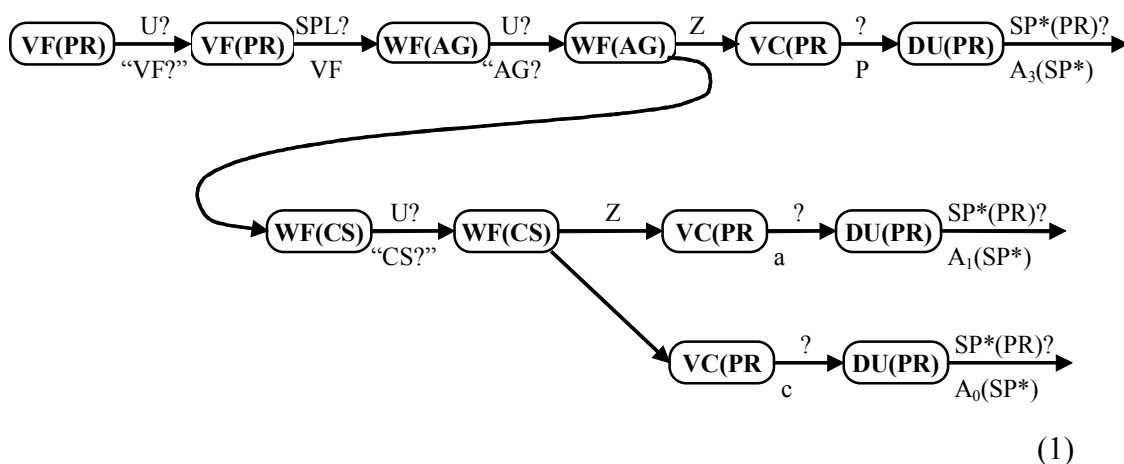
**ABSTRACT.** The paper presents an approach based on network method of language (first of all morphologic) algorithms. The net scheme has experienced certain transfigurations: e.g. a dialogue operator can express the object which is referred to by the system in the current step of process developing using left arc label, the right label must reflect the content of a question on the other hand; as for the knot it includes indication to the position where the expected answer can appear. The “respondent” of a question-answer act might be represented by the inner component of the system (a dictionary, a list of super-paradigms, the information gathered during the previous steps) of outer partner, user (U). Exactly the latter one ensures the interactive character of the system. An example of building the units system operating simultaneously is given in the paper. The aspect of such approach is demonstrated by a simple, though important, example. In particular, the initial of sentence synthesis which generates the central (core) structure of an expression. The section of the process of synthesis provides the core of the central structure, the formation of the verb and its immediate constituents (agent, object, etc.). Building of a system, in its turn, is based on the notion of the verb super-paradigm, which implies the combination of verb paradigms which are derived from one and the same lexeme and addresses the same group of semantic roles, though with a variety of means. How the information that is sufficient for morphologic forming of each member is gathered in the process of dialogue with the user is shown with a particular example. In the last section there is considered applied as well as the fundamental importance of interactive approach: on the one hand, it simplifies and improves the most complicated system, such as automatic translation, on the other hand, the interactive model of a language can be considered a primary imitation of relation between thinking represented by outer partner, user (U) and language – by such organised synthetic system. © 2013 Bull. Georg. Natl. Acad. Sci.

**Key words:** *interactive algorithm, sentence synthesis, morphologic nets.*

The use of the operation system period in the preceding section is illustrated by the scheme (1-5) given below. This scheme produces in the course of interactive regime the main component of Georgian simple sentence, its core structure including the verb

(“core”) and its main actants.

The process realized by this scheme is subject to the following basic restrictions: 1) it deals with verbs belonging to “regular” super-paradigm (SD) only; 2) it takes into account the verb forms of narrative mood



only. As the result of the second restriction the set of rows (mc'k'rivi), to which may belong the produced verb, includes five members only: pr(esent), f(uture), p(ast), pp (past perfect) and rp (resultative perfect). The first trio of this symbolic (pr, f, p) is identical with designation of the tense (T) category values.

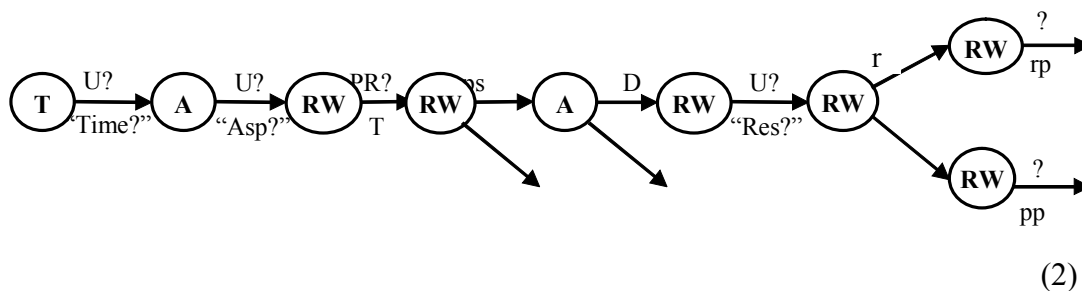
Each step of the (1-5) process implies some definite meaningful purpose. So, the first of them (1) receives from the user (U) the initial word forms, which "point" at the paradigms of the final verb and actant forms. After that (1) procedure defines the value of the voice category (VC) on the basis of semantic roles (SRi) combination presented by the user; and this information, in its turn, is sufficient for definition of concrete paradigm, to which the output verb form VF (DR) should belong.

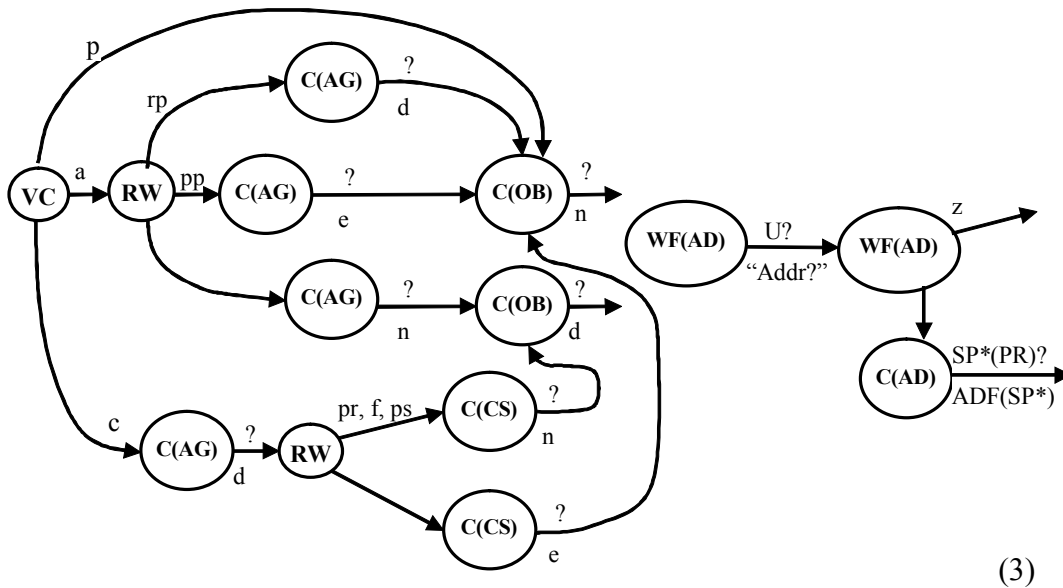
The field for search of VF becomes narrower after the implication of (2), which defines the row (RW) of this paradigm, which includes the output VF. This definition is based on the values of these tense (T) and aspect (A) supplied by user (U).

The combination of voice (VC) and row (RW) values is sufficient for the definition of case (C) values characterizing the actant expression of semantic roles (SRi), and the procedure (3) serves just this purpose.

The next step (4) completes the data necessary for the choice of word forms (WF) expressing all semantic roles belonging to the core structure, that is, it receives number (N) values from the user and in addition defines the value of person category (PS) marking all SRi (immediately from corresponding dictionary units (DU)). The sum of C (case), N (number) data supplies the possibility for the choice of the correct output WF among the members of the paradigm generated by PG of DU corresponding to the initial WF proposed by the user.

The last step (5), firstly, singles out the semantic roles (SRi), which should be explicitly addressed by VF affixes (on the basis of VC and RW values), and then sends their N and PS values together with RW value to the paradigm generator (PG) of (already de-





defined on the (1) step dictionary unit (DU).

As the final result, this interactive system (1-5) produces correct expression (VF, WF) for all members of core structure: PR and SRi.

Perhaps the simplest example will make more transparent the functioning of this system.

Suppose that the U answers the question about the process, which takes place, by the input xat'va ('painting') and then (by means of corresponding questions) "acknowledges" as an agent of the process - k'aci ('man') and as its object (OB) - surati ('picture'), "ignoring at the same time the causer (CS). The AG, OB pair (without CS) points at the active (a) paradigm of corresponding SP, defined by initial input (xat'va). After this U makes more exact VF defining tense (T) and aspect categories: past (p) and perfect (p), which is sufficient for the choice of the row (RW) - pp.

After this U must define the number of categories (N) of AG, OB only (the person values (PS) will be

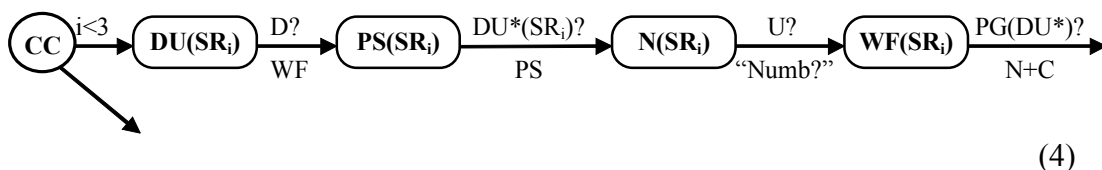
received from the dictionary): N(AG)=N(OB)=3, PS(AG)=PS(OB)=3. Meantime (3) will define the case values of both SRi: C(AG)=e, C(OB)=n, and in conclusion the system will produce:

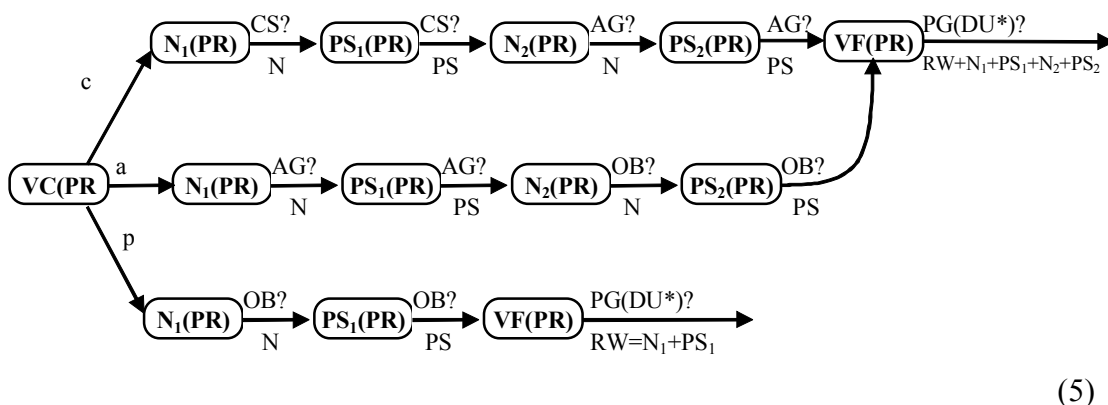
k'acmadaxat'asurati ('the man has painted a picture')

processing from the input sequence: xat'va + k'aci + surati.

**Conclusion.** In conclusion we give some considerations about the main component of the system defining its peculiarities: that is about the dialogue between user and system. Style and content of this dialogue exerts an essential influence on the possibilities of the system applications and the level of its use effectiveness.

At the same time the correct choice of these contact means radically depends on the kind of the user itself. Of course, the simplest solution of this problem takes place when the human partner of dialogue is represented by some "expert", that is by a person





not only speaking the given language, but also having (some) knowledge of its theory (in the first instance, - of its grammar). Such user will understand the questions formulated in the usual linguistic terms without any problems and give them correct answers quite acceptable for the system. Nevertheless such kind of interactivity may be quite useful, and even - necessary, for testing and correction of the system.

After this, a duly tested and corrected system can serve as a source of input for synthesizing the direction of model functioning.

The main theoretical value of this “input source” is its intermediate position between speaker’s intension and its verbal expression, between “thought”, as the “soul” of utterance, and its language form, its “body” (according to Hegel). It can be imagined on, the most primitive (“naïve”) level of abstraction that transformation from “thought” to utterance begins some analogue of “pattern recognition”, where the part of “patterns” is played by the lexical units, and then precedes by the structuring of these units in

correct grammatical sentence. Obviously, the proposed system acts alike to this primitive scheme and thus may serve as some “computer support” for its further development and perfection.

The important aspect of such psycho-linguistic researches is addressed to “ordinary” (not “expert”) persons as objects of these tests. This condition, in its turn, requires a simpler, more usual, and style of communication with a computer system understandable for them. So, in the case of the above mentioned example more felicitous should be questions as: “What is going on?”, “What happens?” (xat’va - ‘painting’); “what is painted” (surati - ‘picture’); “Who paints?” (k’aci - ‘man’); “Does he paint or has finished already?” (xat’avs - ‘is painting’ - T=pr); ...

In conclusion two more fields of possible application for this system can be mentioned. On the ground of some further development it may be used as a basis for analysis of text and in the frames of automatic translation system: in both cases the already presented text will play the part of user.

ენათმეცნიერება

## ქსელურ სისტემაზე დაფუძნებული ენის ინტერაქტიული მოდელის სტრუქტურა

გ. ჩიკოიძე, ა. ჩუტკერაშვილი, ნ. ჯავაშვილი

საქართველოს ტექნიკური უნივერსიტეტის ა. ელიაშვილის მართვის სისტემების ინსტიტუტი

(წარმოდგენილია აკადემიკოს ა. არაბულის მიერ)

ნაშრომში გრძელდება ინტერაქტიული ენობრივი მოდელის წარმოდგენის თემა და შემოთავაზებულია ენობრივი (პირველ რიგში – მორფოლოგიური) ალგორითმების ქსელურ მეთოდზე დაფუძნებული მიდგომა. ქსელურმა სქემამ ინტერაქტიული ასპექტის გასათვალისწინებლად გარკვეული ტრანსფორმაცია განიცადა: მაგალითად, “დილოგური” ოპერატორი რეალის მარცხენა ჭდის საშუალებით გამოხატავს ობიექტს, რომელსაც სისტემა შეკითხვით მიმართავს პროცესის განვითარების მიმდინარე ბიჯზე: ამჟღავნებდა მარჯვენა ჭდემ უნდა ასახოს შეკითხვის შინაარსი, კვანძი კი მოიცავს მითითებას პოზიციასზე, სადაც მოსალოდნელი პასუხი უნდა დაფიქსირდეს. კითხვა-პასუხის აქტის “რესპონდენტს” შეიძლება წარმოადგენდეს სისტემის შიდა-კომპონენტი (ლექსიკონი, სუპერ-პარადიგმათა სია, წინა ეტაპზე დაგროვილი ინფორმაცია) ან გარე “პარტნიორი”, ანუ მომხმარებელი (U). სწორედ ეს უკანასკნელი ვარიანტი უზრუნველყოფს სისტემის ინტერაქტიულობის თვისებას. წინა ნაშრომი შეიცავდა ზემოხსენებული მიდგომის ცალკეული კომპონენტების (ოპერატორების და მათი არგუმენტების შესატყვისი ინფორმაციული სტრუქტურების) აღწერას. ამ სტატიის კი მოცემულია ელემენტებისაგან ერთობლივად მოქმედი სისტემის აგების მაგალითი. მიდგომის ეს ასპექტი დემონსტრირებულია მარტივ, მაგრამ მეტად მნიშვნელოვან ნიმუშებზე. სახელდობრ, წინადადების სინთეზის საწყის ეტაპზე, რომელზეც ხდება გამონათქვამის ცენტრალური (Core) სტრუქტურის წარმოქმნა. სინთეზირების პროცესის ეს მონაკვეთი უზრუნველყოფს ცენტრალური სტრუქტურის ბირთვს, ანუ ზმნის და მისი უშუალო აქტანტების (აგენტის, ობიექტის და ა.შ.) ფორმირებას. სისტემის აგება კი, თავის მხრივ, ზმნური სუპერ-პარადიგმის ცნებას ეფუძნება, რომელიც გულისხმობს ზმნური პარადიგმების ერთობლიობას, რომლებიც ერთი და იმავე ლექსემისაგანაა ნაწარმოები და სემანტიკური როლების ერთსა და იმავე ჯგუფს მიმართავენ, თუმცა სხვადასხვა საშუალებებით. კონკრეტულ მაგალითზე ნაჩვენებია, მომხმარებელთან დიალოგის მსვლელობისას როგორ ხდება ინფორმაციის დაგროვება, რომელიც საკმარისია ცენტრალური სტრუქტურის ყველა წევრის მორფოლოგიური გაფორმებისთვის. ნაშრომის დასკვნით ნაწილში აღნიშნულია ინტერაქტიული მიდგომის როგორც გამოყენებითი, ისე ფუნდამენტური მნიშვნელობა: ერთი მხრივ, ის გაამარტყვებს და სრულყოფილს გახდის ისეთ ურთულეს სისტემებს, როგორცაა ავტომატური თარგმანი; მეორე მხრივ კი, ენის ინტერაქტიული მოდელი შეიძლება განხილული იქნეს, როგორც პირველადი იმპლაცია მიმართების აზროვნებასა და ენას შორის, სადაც აზროვნებას წარმოადგენს გარე პარტნიორი, მომხმარებელი (U), ენას კი – თავად ასე ორგანიზებული სინთეზური სისტემა.

### REFERENCES:

1. G. Chikoidze, A. Chutkerashvili, N. Javashvili (2013), Bull.Georg.Natl.Acad.Sci., 7, 1: 119-126.
2. G. Chikoidze (2004), Net Representation of Morphologic Processors. Monograph, Tbilisi (in Russian).
3. G. Chikoidze (2012), Bull.Georg.Natl.Acad.Sci. 6, 3:145-153 Tbilisi.

Received june, 2013