ITHEA

International Journal

INFORMATION MODELS & ANALYSES

2012 Volume 1 Number 1

International Journal
INFORMATION MODELS & ANALYSES
Volume 1 / 2012, Number 1

IJ IMA is official publisher of the scientific papers of the members of the ITHEA® International Scientific Society

IJ IMA rules for preparing the manuscripts are compulsory.
The **rules for the papers** for ITHEA International Journals as well as the **subscription fees** are given on www.ithea.org .
The camera-ready copy of the paper should be received by ITHEA® Submission system http://ij.ithea.org .
Responsibility for papers published in IJ IMA belongs to authors.

General Sponsor of IJ IMA is the **Consortium FOI Bulgaria** (www.foibg.com).

International Journal "INFORMATION MODELS AND ANALYSES" Vol.1, Number 1, 2012

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with
Institute of Mathematics and Informatics, BAS, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Universidad Politechnika de Madrid, Spain,
Hasselt University, Belgium
Institute of Informatics Problems of the RAS, Russia,
St. Petersburg Institute of Informatics, RAS, Russia
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia,
and Federation of the Scientific - Engineering Unions /FNTS/ (Bulgaria).

Publisher: **ITHEA®**
Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: info@foibg.com
Technical editor: Ina Markova

Printed in Bulgaria

ISSN 1314-6416 (printed) **ISSN 1314-6424 (CD)** **ISSN 1314-6432 (Online)**

PREFACE

The ITHEA International Journal ***"Information Models and Analyses"*** (IJ IMA) is aimed to present advances in the information modeling theory and practice.

The concept "Information Model" is fundamental for the Informatics. There exist many approaches to define this concept as well as a long list of names of scientists who worked to define more precisely the concept "Model", and respectively, the "Information Model". The list contains the names of N.Wiener and A.Rosenblueth, V.M.Glushkov, M.W.Wartofsky and many others.

In the beginning, the concept "Information model" has been used to denote mainly one of the activities of using the computer technique. Maybe, it is the most popular understanding. Nevertheless, the definition of the concept "information model" may and needs to be extended to cover the new scientific paradigms, which come from the Modern Informatics.

Now, it is time to think more general about information models and corresponded analyses, to think more general about the "information modeling" as theory and practice.

Sofia, 24.05.2012.

Krassimir Markov

MULTIPLE-MODEL DESCRIPTION AND STRUCTURE DYNAMICS ANALYSIS OF ACTIVE MOVING OBJECTS CONTROL SYSTEM

Boris Sokolov, Rafael Yusupov, Vyacheslav Zelentsov

Abstract: *Proposed developed concept of active moving object (AMO) was used to form statement of control problems which was investigated. Depending on the type of AMO can move and interact in space, air, on the ground, in water. In this paper AMO was interpreted as space-facilities (SF). The unified description of various control processes lets synthesize simultaneously both technical and functional structures of SF control system (CS). It is important that the presented approach extends new scientific and practical results obtained in the modern control theory to the field of space programs.*

Keywords: *active moving object, optimal program control, multi-criteria multiple-model description applied area.*

ACM Classification Keywords: *D.2.11 Software Architectures and F.1.1 Models of Computation.*

Introduction

One of the main features of modern complex technical systems (CTS), particularly orbital and ground-based space facilities, is the changeability of their parameters and structures caused by objective and subjective reasons at different stages of the CTS life cycle. In other words, we always come across the CTS structure dynamics in practice (Fig. 1). In Fig. 1 S^j is a number of CTS structural state. Under the existing conditions to increase (stabilize) CTS work potentialities and capacity a structure control (including the control of CTS structure reconfiguration) is to be performed.

According to the specifics of the structure-dynamics control problems, they belong to the class of the CTS structure-functional synthesis problems and the problems of program construction for CTS development.

The main disadvantage of the problems belonging to the above class is that, optimal control programs for CTS main elements and subsystems can be implemented only when the lists of functions and algorithms for control and information processing in these subsystems and elements are known.

The distribution of the functions and algorithms among the CTS elements and subsystems, in its turn depends upon the control laws actual for these elements and subsystems.

The described contradictory situation is complicated by the changes of CTS parameters and structures caused by different factors during the CTS life cycle.

Currently, the class of problems being reviewed is not examined thoroughly enough. New theoretical and practical results were obtained at the following directions of investigations:

- the synthesis of the CTS technical structure for the known laws of CTS functioning (the first direction) [Casti, 1979], [Klir, 1985], [Zvirkun, Akinfiev, Filippov, 1985], [Zvirkun, Akinfiev, 1993], [Zvirkun, 1982];

- the synthesis of programs for CTS construction and development without taking into account the periods of parallel functioning of the existing and new CTS (the third direction) [Tsurlov, 1989], [Zvirkun, Akinfiev, 1993];
- the parallel synthesis of the CTS technical and functional structures (the forth direction) [Sokolov, 1999], [Tsurlov, 1989], [Zvirkun, Akinfiev, 1993].

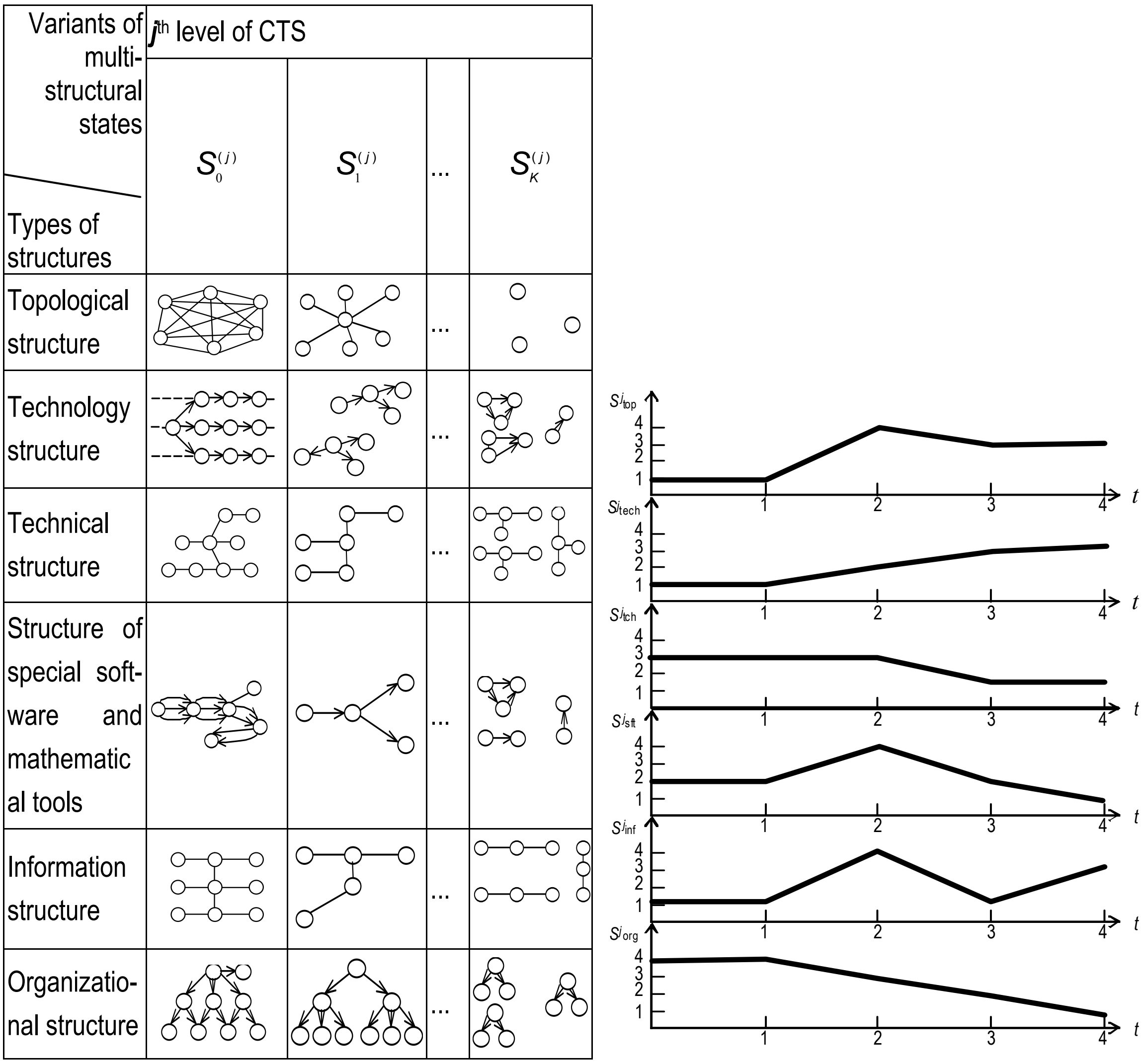


Fig. 1. Diagrams of CTS structure dynamics Possible variants of CTS structure dynamics.

Let us briefly consider the main results and state-of-the-art along the above mentioned directions of investigations.

A great deal of work regarding various problems of the CTS technical structure synthesis (the first direction of investigations) is accomplished in Russia and abroad [Casti, 1979], [Klir, 1985], [Zvirkun, Akinfiev, Filippov, 1985], [Zvirkun, Akinfiev, 1993], [Zvirkun, 1982].

The synthesis (selection) of a CTS structure (structures) was usually reduced to the following general optimization problem [Zvirkun, Akinfiev, Filippov, 1985], [Zvirkun, 1982]:

$$\bar{S}\left\{\left[\bar{f} \subset \bar{F}(\bar{\pi})\right]\bar{R}\left[\bar{m} \subset \bar{M}\right]\right\} \to extr\,, \tag{1}$$

$$\bar{\pi} \subset \bar{P}\,, \tag{2}$$

$$\bar{f} \subset \bar{F}(\bar{\pi})\,, \tag{3}$$

$$\bar{m} \subset \bar{M}\,, \tag{4}$$

where $\bar{P}$ is a set of feasible control principles (algorithms); $\bar{F}$ is a set of interrelated functions (tasks, operations) that can be performed by the system. For each subset $\bar{\pi} \subset \bar{P}$ there exists the set $\bar{F}(\bar{\pi})$ the realizations sufficient for the given principles (algorithms) should be chosen from, i.e., it is necessary to choose $\bar{f} \subset \bar{F}(\bar{\pi})$; $\bar{M}$ is a set of CTS possible elements such as information processing and transceiving facilities, control units, service terminals, etc; the map $\bar{R}$ takes $\bar{F}$ to $\bar{M}$.

It is stated that the optimal map $\bar{F}$ returns an extremum to some objective function (functions) $\bar{S}$ under given constraints.

The modifications of the considered problem concern the aspects of uncertainty and multi-criteria decision-making.

The complexity of the synthesis problem (1)–(4) is mainly caused by its high dimensionality that is by the number of variables and constraints in the detailed problem statement. That is why the methods of decomposition, aggregation and sub-problem coordination are widely used. Another peculiarity complicating the problem is the integer-valued variables.

The features of the structure synthesis problem were thoroughly considered in the works [Zvirkun, Akinfiev, Filippov, 1985], [Zvirkun, Akinfiev, 1993], [Zvirkun, 1982]. The authors proposed a multi-level interrelated complex of analytical and simulation models based on decomposition and aggregation approach.

Studies on structure synthesis problems (1)–(4) confirm [Zvirkun, Akinfiev, Filippov, 1985], [Zvirkun, Akinfiev, 1993], [Zvirkun, 1982] that when CTS elements and subsystems cannot manage peak data traffic, then the law of elements functioning ought to be optimized (the second direction of investigations).

The laws and algorithms of hierarchical system functioning, as well as problems of functional synthesis have been investigated in Russia and abroad for more than 40 years within the developing control theory [Athaus, Falb, 1966], [Bryson, Yo-Chi, 1969], [Moiseev, 1974], [Pontriagin, etc., 1961], [Singh, Titli, 1978], [Tsypkin, 1971], [Vasil'ev, 2001]. Therefore, it is reasonable to consider the particular scope of these investigations in accordance with the aims of this paper. Here we discuss the problems of CTS structure-dynamics control.

Fig. 2 from [Doganovskii, Oseranii, 1990] shows the classification of CTS, the concept of structure dynamics control was applied to. The numbers denote the following classes of the systems: 1 — CTS with controllable structure dynamics; 2 — basic reconfigurable CTS; 3 — systems with coordinate-parametric control (SCPC); 4 — systems with active controllable technologies (SACT); 5 — integrated active-control systems (IACS); 6 — systems of alternative control and multiple-mode control; 7 — systems of fault-tolerant self-recovering control; 8 — systems of intellectual control.

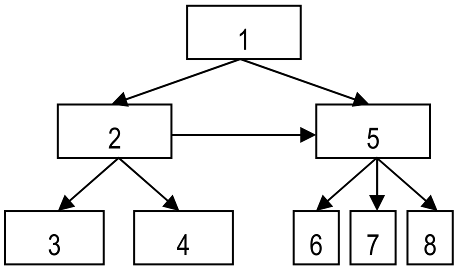


Fig. 2. Classification diagram of reconfigurable systems.

The control problems for basic reconfigurable CTS were the best investigated ones. Interesting fundamental and practical results were obtained in this field [Bryson, Yo-Chi, 1969], [Napolitano, Swaim, 1989], [Ohtilev, Sokolov, Yusupov, 2006], [Van der Velde, 1984].

The investigations towards creation and application of integrated active-control systems are still at the initial phase, especially for the systems of controllable structure dynamics with intellectual control elements. These systems are functioning under mutable objectives and external perturbation actions that can be purposeful and/or not purposeful [Fleming, Richel, 1975], [Gupta, Sinka, 1996], [Ohtilev, Sokolov, Yusupov, 2006], [Russell, Norvig, 1995], [Shannon, 1975], [Oreu, Zeigler, Elzas, 1984], [Sokolov, 1999], [Tsypkin, 1971], [Vasil'ev, 2001], [Yusupov, Rozenwasser, 1999].

The growth of CTS complexity along with the increasing importance of uncertainty factors at all stages of CTS functioning necessitates new approaches for control-system construction.

The most perspective approach, namely, intellectual control has arisen within artificial-intelligence investigations [Gupta, Sinka, 1996], [Russell, Norvig, 1995], [Vasil'ev, 2001].

Fig. 3 presents the multilevel scheme of CTS control.

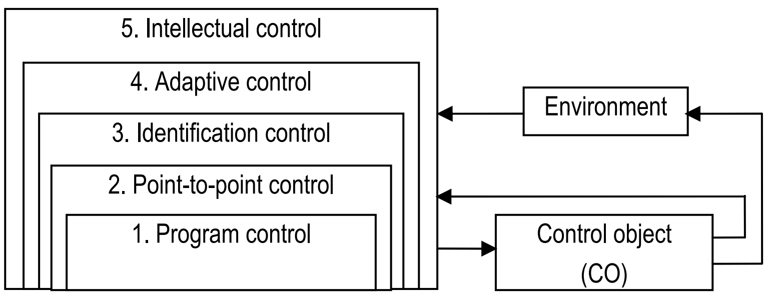


Fig. 3. Multilevel structure of CTS control processes.

Fig. 4, 5 from [Vasil'ev, 2001] show respectively the sources of intellectual control and the relations of scientific disciplines forming the theory of intellectual control. The work [Vasil'ev, 2001] gave a detailed analysis of intellectual-control investigations that have been carried out in Russia and abroad during the last 30 years. The most interesting recent scientific and practical results in the field of CTS control were received on the basis of production languages with fuzzy rules and by means of neural networks [Gupta, Sinka, 1996], [Russell, Norvig, 1995], [Vasil'ev, 2001].

To accomplish the analysis of state-of-the-art CTS structure-dynamics control let us briefly consider the third and the forth directions of structure synthesis that were mentioned at the beginning of the section.

There are several works [Casti, 1979], [Klir, 1985], [Zvirkun, Akinfiev, 1993] related to theoretical basics for the synthesis of CTS development programs. Several iterative procedures for particular structure-functional synthesis problems concerning the early stages of CTS life cycle are currently known. Nevertheless, the obtained results do not conform to dynamics of the environment at the CTS operating stage when the time factor is rather important [Sokolov, 1999]. As well, the peculiarities of distributed CTS are not considered thoroughly enough for the operation phase.

Therefore, the development of new theoretical bases for CTS structure-dynamics control is very important now. From our point of view, the theory of structure-dynamics control will be interdisciplinary and will accumulate the results of classical control theory, operations research, artificial intelligence, systems theory, and systems analysis. These ideas are summarized in Fig. 6 that is a modification of Fig. 5 for the subject of this paper.

Here, as the first step to the new theory, we introduce conceptual and formal description of CTS structure dynamics. The interpretation of the constructed models concerns SF control systems.

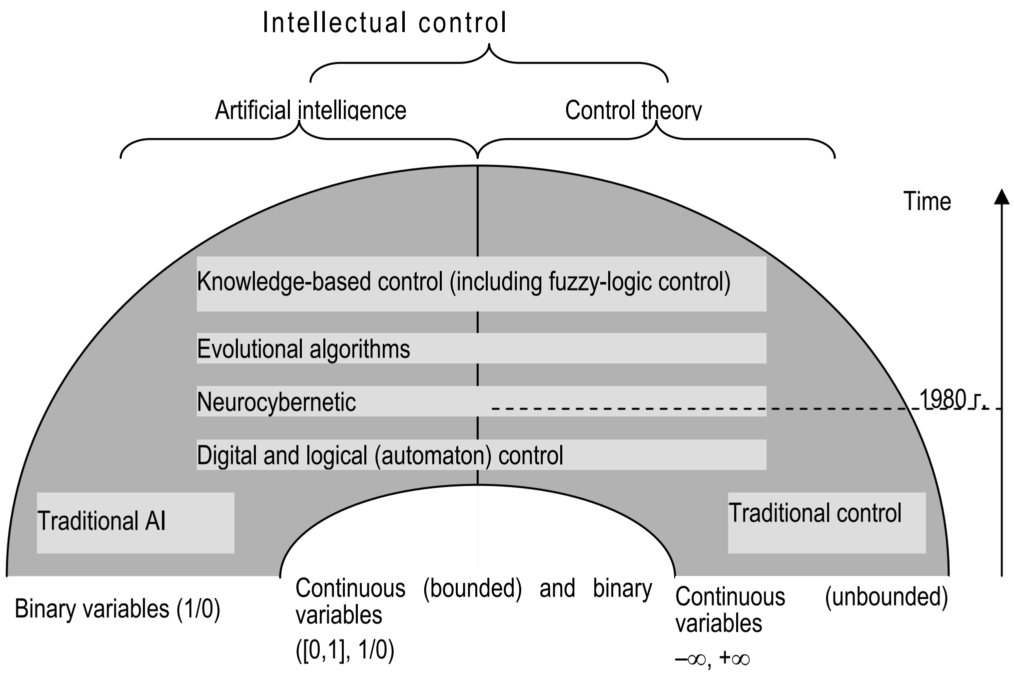


Fig. 4. Two main sources of intelligent control.

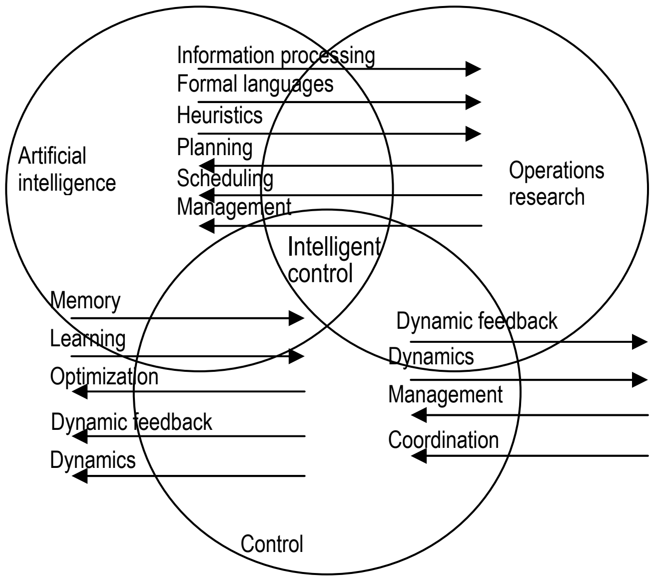


Fig. 5. Intellectual control as a scope of investigations.

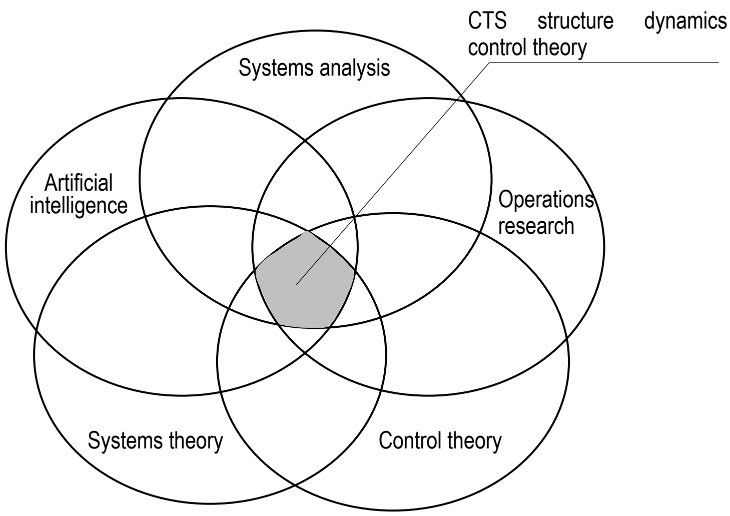


Fig. 6. The theory of structure-dynamics control as a scope of interdisciplinary investigations.

The complex of SF program-control models is discussed in this paper. These models give unified technology for an analysis and optimization of various processes concerning spacecraft information receiving, transferring and processing. The main advantage of the constructed models is that the structural and functional constraints for SF control process are defined explicitly. By now, the prototype programs realizing models were developed and used for evaluation of SF CS abilities and for planning of SF CS operation.

II. CONCEPTUAL MODEL FOR CONTROL PROCESSES OF ORBITAL AND GROUND-BASED SPACE FACILITIES

Under the orbital space facilities (OSF) a group of spacecrafts having a common mission is meant. The ground-based space facilities (GSF) are hardware-software complexes supporting spacecraft functioning at the phase of orbital flight. These complexes are the parts of ground-based control systems (GCS) such as control centers (CC) for spacecrafts of different missions, control stations (CST), OSF-interacting stations (IS), central control station (CCS), and telecommunication system (TS) providing for communication between enumerated elements. The groups of OSF and GSF form the SF control system (SF CS).

In Fig. 7 the generalized structure of SF CS is shown. The following notation was used: HSC is a unified hardware-software complex for control automation; the symbols ○ denote spacecrafts of some group; the symbols ⊗ denote objects-in-service (OS). The last are sources or recipients of information being processed, they can be moving or stationary, and in particular, they can represent some water, ground, or space areas that are interesting for the recipients. OS can change their state during the information, material, or energy interactions with OSF.

The preliminary investigations confirm that the most convenient concept for the formalization of SF control processes is the concept of an active moving object (AMO). In general case, it is an artificial object (a complex of devices) moving in space and interacting (by means of information, energy, or material flows) with other AMO and OS [Sokolov, 1999], [Sokolov, Kalinin, 1995].

Fig. 8 shows general structure of AMO as an object being controlled. It is seen that AMO consists of four subsystems relating to four processes (functioning forms): moving, interaction with OS and other AMO, functioning of the main (goal-oriented) and auxiliary facilities, resources consumption (replenishment).

The four functions of AMO are quite different, though the joint execution of these functions, the interaction being the main one, provide for AMO new characteristics. Thus, it becomes a specific object of investigation, and AMO control problems are strictly different than classical problems of mechanical-motion control. The proposed structure of AMO can be widely interpreted. It gives a common basis for description of OSF, GSF, and OS.

To construct the models of AMO (SF or OS) control we should firstly formulate the goal of its functioning. This goal is to be related to the interaction between other AMO and OS. Secondly, the sequence of operations that lead to the specified goal should be determined. Each operation is characterized by its parameters. Some of parameters describe the results of the operation performance (amount of work, quality, duration, resource consumption, etc), the others present the characteristics of material or information flows necessary to perform the operation.

The conceptual model of SF control can be presented by means of state (macro-state) diagrams. Fig. 9 shows the fragment of a diagram for transition from SF general states. The following notations were used: 1 — execution of SF goal tasks; 2 — reserve state of SF; 3 — SF technical service and repair; 4 — SF motion for the goal tasks; 5 — SF motion after goal-tasks execution. In [Sokolov, 1999], [Sokolov, Kalinin, 1995] other examples of diagrams are presented.

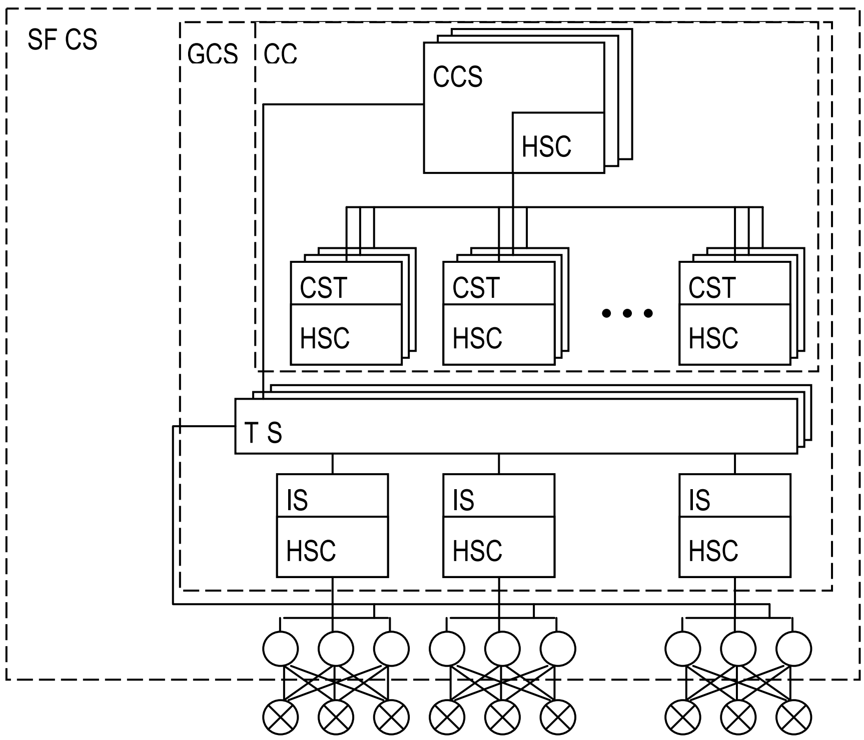


Fig. 7. The generalized technical and organizational structure of SF CS.

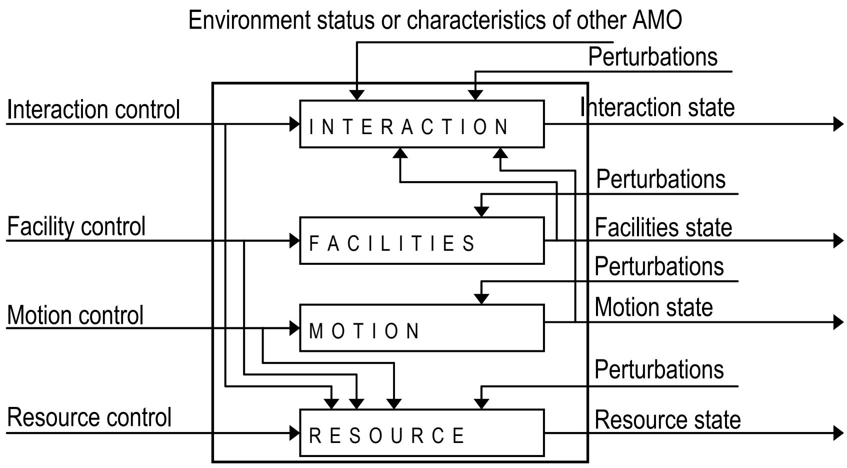


Fig. 8. General block diagram of AMO.

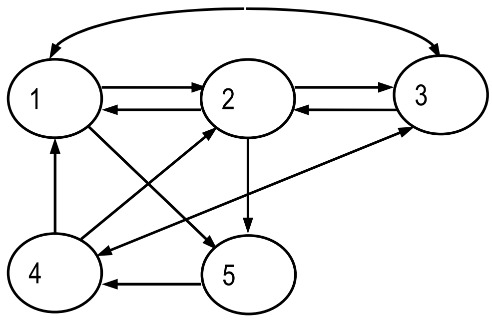


Fig. 9. The fragment of general-state transition diagram.

Thus, at the conceptual level, the process of SF functioning can be described as a process of operation execution, while each operation can be regarded as a transition from one state to another one. Meanwhile, it is convenient to characterize the SF state by the parameters of operations.

The particular control models are based on the dynamic interpretation of operations and the previously developed particular dynamic models of AMO functioning.

III. Multiple-model description of space-facilities control processes

A. Set-theory based model

In accordance with the proposed conceptual model of SF control, let us introduce the following basic sets and structures:

$B = \{B_i, i \in M = \{1,\ldots,m\}\}$ is a set of objects (subsystems, elements) that are embodied in SF CS and are necessary for its functioning;

$\bar{B} = \{\bar{B}_i, i \in \bar{M} = \{1,\ldots,\bar{m}\}\}$ is a set of external objects, (subsystems, elements) interacting with SF CS (the interaction may be informational, energy or material);

$\tilde{B} = B \cup \bar{B}$ is a set of the considered objects;

$\tilde{C} = C \cup \bar{C} = \{C_1, C_2, \ldots, C_m\} \cup \{\bar{C}_1, \bar{C}_2, \ldots, \bar{C}_{\bar{m}}\}$ is a set of channels that are used by SF and OS for interaction;

$C = \{C_\lambda^{(i)}, \lambda \in \Lambda_i, i \in M\}$ is a set of channels (hardware facilities) that exist on the considered SF;

$D = \left\{D^{(c)} \cup D^{(i)}, i \in M\right\}$ is a set of SF CS operations;

$D^{(i)} = \left\{ D_\alpha^{(i)}, \alpha \in K_i^{(o)} = \{1,\ldots,s_i\} \right\}$ is a set of interaction operations with the object B_i;

$D^{(c)} = \left\{\{D_{iwf\eta_1}^{(c,1)}\} \cup \{D_{iwf\eta_1}^{(c,2)}\} \cup \{D_{iwf}^{(c,3)}\}, i \in M, w \in NW^{(i)}, f \in NF^{(w)}, \eta_1 \in NH_1\right\}$ is a set of SF CS macro-operations;

$\{D_{iwf\eta_1}^{(c,1)}\}$ is a set of macro-operations for the object B_i and its macro-state S_{iwf} at the control cycle η_1; the sets of subscripts $M = \{1,\ldots,m\}$, $NW^{(i)} = \{1,\ldots,K_W^{(i)}\}$, $NW^{(w)} = \{1,\ldots,K_F^{(w)}\}$, $NH_1 = \{1,\ldots,E_H\}$ are respectively used for the macro-states of the object B_i, for the place numbers of objects B_i in the macro-state «*w*», for the control cycles of the object B_i;

$\{D^{(c,2)}_{iwf\eta_1}\}$ is a set of auxiliary operations;

$\{D^{(c,3)}_{iwf}\}$ is a set of macro-operations for B_i transitions from the macro-state $S_{iw'f''}$ to the macro-state S_{iwf};

$\Phi = \left\{ \{\Phi S^{(i)}_{\pi}\} \cup \{\Phi N^{(i)}_{\mu}\}, i \in M, \pi \in K^{(p,1)}_i = \{1,\ldots,k^{(p,1)}_i\}, \mu \in K^{(p,2)}_i = \{1,\ldots,k^{(p,1)}_i\} \right\}$ is a set of SF CS resources;

$\Phi S^{(i)} = \left\{ \Phi S^{(i)}_{\pi}, \pi \in K^{(p,1)}_i \right\}$ is a set of non-storable resources of the object B_i;

$\Phi N^{(i)} = \left\{ \Phi N^{(i)}_{\mu}, \mu \in K^{(p,2)}_i \right\}$ is a set of storable resources of the object B_i;

$P = \left\{ \{P^{(i)}_{<\alpha',\rho>}\} \cup \{P^{(i,j)}_{<\alpha,\rho>}\}, i \in M, \alpha' \in K^{(o)}_i, \nu \in K^{(o)}_{<i,j>}, \rho \in K^{(n)}_i \right\}$ is a set of SF CS flows;

$P^{(i)} = \left\{ \{P^{(i)}_{<\alpha',\rho>}\}, i \in M, \alpha' \in K^{(o)}_i, \rho \in K^{(n)}_i \right\}$ is a set of flows (energy flows, material flows, and information flows) produced by or necessary for the object B_i;

$P^{(i,j)} = \left\{ P^{(i,j)}_{<\alpha,\rho>}, i, j \in M, \rho \in K^{(n)}_i \right\}$ is a set of flows (energy flows, material flows, and information flows) produced when the objects B_i and B_j interact;

$G = \left\{ G_{\chi}, \chi \in NS \right\}$ is a set of SF CS structural types, where the main types of structures are the topologic (spatial) structure, the technology (functional) structure, the technical structure, the structures of mathematical and software tools, and the organizational structure.

To interconnect the structures let us consider the following dynamic alternative multi-graph (DAMG):

$$G^t_{\chi} = \left\langle X^t_{\chi}, F^t_{\chi}, Z^t_{\chi} \right\rangle, \quad (5)$$

where the subscript χ characterizes the SF CS structure type, $\chi \in NS = \{1,2,3,4,5,6\}$ (here 1 indicates the topologic structure, 2 indicates the functional structure, 3 indicates the technical structure, 4 and 5 indicate the structures of mathematical and software tools, 6 indicates the organizational structure, the time point *t* belongs to a given set T; $X^t_{\chi} = X^t_{\chi} = \{x^t_{\chi l}, l \in L_{\chi}\}$ is a set of elements of the structure G^t_{χ} (the set of DAMG vertices) at the time point t; $F^t_{\chi} = \{f^t_{<\chi,l,l'>}, l, l' \in L_{\chi}\}$ is a set of arcs of the DAMG G^t_{χ}; the arcs represent relations between the DAMG elements at time t; $Z^t_{\chi} = \{f^t_{<\chi,l,l'>}, l, l' \in L_{\chi}\}$ is a set of parameters that characterize relations numerically.

The graphs of different types are interdependent, thus, for each particular task of SF CS structure–dynamics control the following maps should be constructed:

$$M^t_{<\chi,\chi'>} : F^t_{\chi} \to F^t_{\chi'}, \quad (6)$$

compositions of the maps can be also used at time *t*:

$$M^t_{<\chi,\chi'>} = M^t_{<\chi,\chi_1>} \circ M^t_{<\chi,\chi_2>} \circ \ldots \circ M^t_{<\chi'',\chi'>}. \quad (7)$$

A multi-structural state can be defined as the following inclusion:

$$S_\delta \subseteq X_1^t \times X_2^t \times X_3^t \times X_4^t \times X_5^t \times X_6^t, \quad \delta = 1,\ldots,K_\Delta . \tag{8}$$

Thus we obtain the set of SF CS multi-structural states:

$$S = \{S_\delta\} = \{S_1,\ldots,S_{K_\Delta}\} . \tag{9}$$

Allowable transitions from one multi-structural state to another one can be expressed by means of the maps:

$$\Pi^t_{<\delta,\delta'>} : S_\delta \to S_{\delta'} . \tag{10}$$

Here we assume that each multi-structural state at time $t \in T$ is defined by a composition (7).

Now the problems of SF CS structure dynamics control can be regarded as a selection of a multi-structural state $S^*_\delta \in \{S_1, S_2,\ldots,S_{K_\Delta}\}$ and of a transition sequence (composition) $\Pi^{t_1}_{<\delta_1,\delta_2>} \circ \Pi^{t_2}_{<\delta_2,\delta_3>} \circ \Pi^{t_f}_{<\delta',\delta>}$ $(t_1,< t_2 < \ldots < t_f)$, under some criterion of effectiveness. The results of the selection can be presented as the optimal program for SF CS structure dynamics control. This program guides the system from a given multi-structural state to the specified one.

Along with the set $\{S_\delta\}$ of multi-structural states let us introduce two additional sets which are necessary for the description of SF CS structure dynamics. We shall use the set $\{S_{\chi\omega}\}$ of structural states (for structures of the type G_χ) and the set $\{S_{iwf}\}$ of macro-states (for the object B_i). Here $\chi \in NS = \{1,\ldots,K_s\}$, $\omega \in N\Omega^{(\chi)} = \{1,\ldots,K^{(\chi)}_\Omega\}\}$, $w \in NW^{(i)} = \{1,\ldots,K^{(i)}_W\}$, $f \in NF^{(w)} = \{1,\ldots,K^{(w)}_F\}$.

All the presented models can be interrelated into a generalized model.

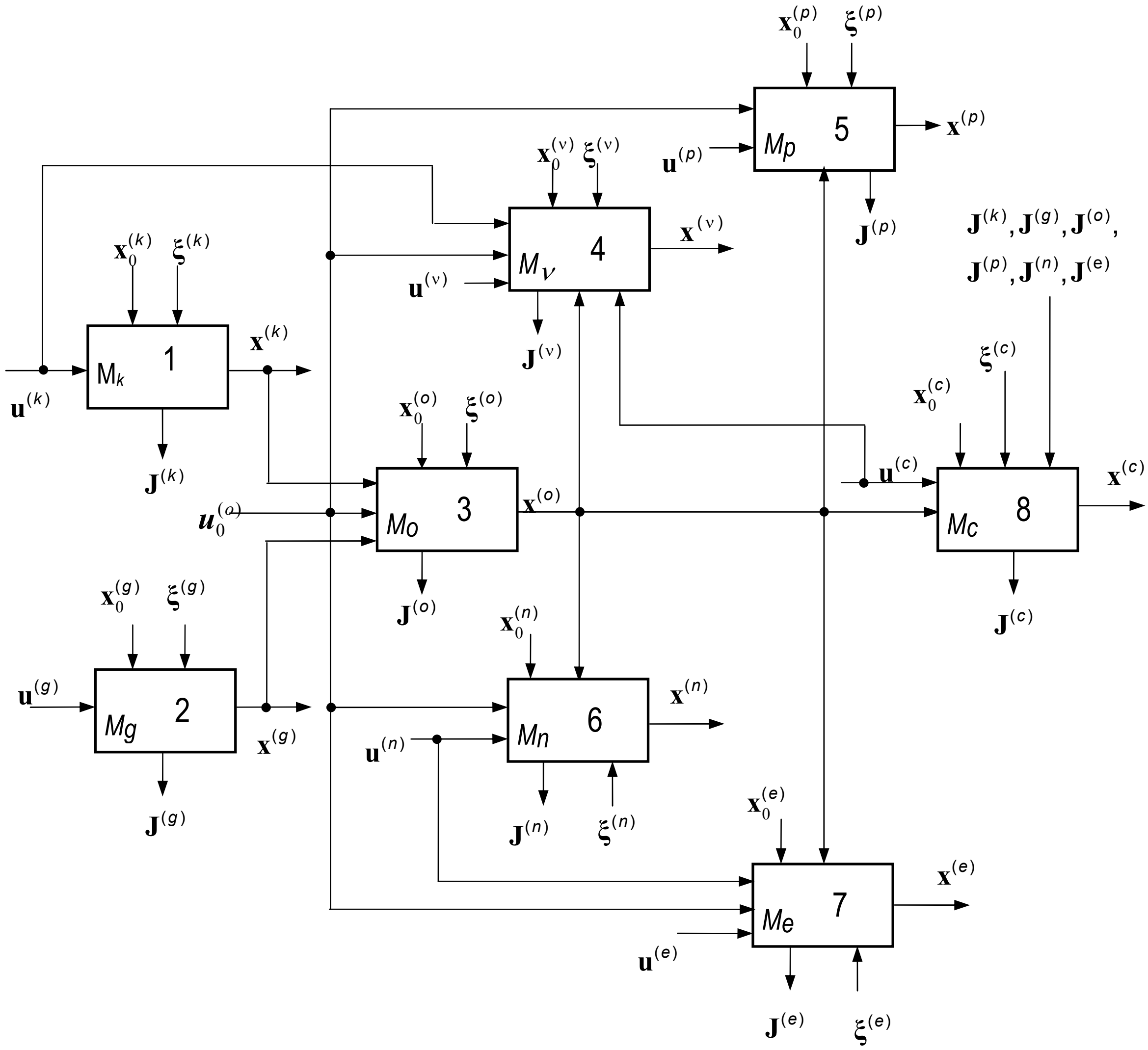
B. Generalized dynamic model of SF CS control processes (*M* model)

$$M = \left\{\mathbf{u}(t) \mid \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},t); \mathbf{h}_0(\mathbf{x}(t_0)) \le \mathbf{O}, \mathbf{h}_1(\mathbf{x}(t_f)) \le \mathbf{O}, \mathbf{q}^{(1)}(\mathbf{x},\mathbf{u}) = \mathbf{O}, \mathbf{q}^{(2)}(\mathbf{x},\mathbf{u}) < \mathbf{O}\right\}, \tag{11}$$

$$J_\theta = J_\theta\left(\mathbf{x}(t),\mathbf{u}(t),t\right) = \varphi_\theta\left(\mathbf{x}(t_f)\right) + \int_{t_0}^{t_f} f_\theta\left(\mathbf{x}(\tau),\mathbf{u}(\tau),\tau\right) d\tau, \ \theta = 1,\ldots,\Theta , \tag{12}$$

where $\mathbf{x} = \| \mathbf{x}^{(g)\mathrm{T}}, \mathbf{x}^{(o)\mathrm{T}}, \mathbf{x}^{(k)\mathrm{T}}, \mathbf{x}^{(p)\mathrm{T}}, \mathbf{x}^{(n)\mathrm{T}}, \mathbf{x}^{(e)\mathrm{T}}, \mathbf{x}^{(c)\mathrm{T}}, \mathbf{x}^{(v)\mathrm{T}} \|^{\mathrm{T}}$ is a vector of SF generalized state, $\mathbf{u} = \| \mathbf{u}^{(g)\mathrm{T}}, \mathbf{u}^{(o)\mathrm{T}}, \mathbf{u}^{(k)\mathrm{T}}, \mathbf{u}^{(p)\mathrm{T}}, \mathbf{u}^{(n)\mathrm{T}}, \mathbf{u}^{(e)\mathrm{T}}, \mathbf{u}^{(v)\mathrm{T}}, \mathbf{u}^{(v)\mathrm{T}} \|^{\mathrm{T}}$ is a vector of generalized control; $\mathbf{h}_0, \mathbf{h}_1$ are known vector-functions that are used for the state $\mathbf{x}$ end conditions at the time points $t = t_0$ and $t = t_f$; the vector-functions $\mathbf{q}^{(1)}, \mathbf{q}^{(2)}$ define the main spatio-temporal, technical and technological conditions for the SF functioning process.

On the whole the constructed model M (11) is a deterministic nonlinear non-stationary finite-dimensional differential system with a reconfigurable structure. Fig. 10 shows the interconnection of models M_g , M_o , M_k , M_p , M_n , M_e , M_c , M_v *, embodied in the generalized model. In Fig. 10 the additional vector-function* $\xi = \| \xi^{(g)\mathrm{T}}, \xi^{(o)\mathrm{T}}, \xi^{(k)\mathrm{T}}, \xi^{(p)\mathrm{T}}, \xi^{(n)\mathrm{T}}, \xi^{(e)\mathrm{T}}, \xi^{(c)\mathrm{T}}, \xi^{(v)\mathrm{T}} \|^{\mathrm{T}}$ *of perturbation actions is introduced. This function describes the impact of the environment upon the SF functioning.*



M_g — dynamic model of CTS motion control; M_k — dynamic model of CTS channel control;

M_o — dynamic model of CTS operations control; M_n — dynamic model of CTS flow control;

M_p — dynamic model of CTS resource control; M_e — dynamic model of CTS operation parameters control;

M_c — dynamic model of CTS structure dynamic control; M_ν — dynamic model of CTS auxiliary operation control

Fig. 10. The scheme of model interconnection.

IV. Methods and algorithms for evaluation of SF CS information-technological and goal abilities

One of the important problems in SF CS structure dynamic control is the evaluation of goal abilities, i.e., potential of the system to perform its missions in different situations. Thus, the preliminary analysis of information and technological and goal abilities (GA and ITA) of SF CS can be used to obtain reasonable means of the objects $B_j, j = 1,..,m$ exploitation under different conditions.

The numerical estimations of SF CS GA and ITA should be based on the system of measures. These measures can be regarded as characteristics of SF CS potential effectiveness. The GA measures characterizing different levels of SF CS are interrelated and have a hierarchical structure.

The leading role of information and technological aspects of the goal-abilities (GA) evaluation is a result of the influence of the technology structure (the structure of SF CS control technology) upon the other SF CS structures (organizational structure, technical structure, etc.) [Casti, 1979], [Klir, 1985], [Moiseev, 1974], [Roy, 1996], [Sokolov, 1999].

So the information and technological abilities (ITA) of a system ought to be evaluated first of all. These abilities can be measured as SF CS capacities [Sokolov, 1999].

The following measures are to be evaluated: the total number of objects in a given macro-state over a fixed time period or at a fixed point of time (this characteristics can be obtained with the model M_c); the total number of working operations performed over a fixed time period $\sigma = (t_0, t_f)$ or by the time point t.

Parallel with the enumerated measures of ITA the following measures of GA can be used: the total possible number of objects-in-service (OS) over the time period σ; the total time that is necessary for the execution of all interaction operations with OS. If the uncertainty factors are considered (the stochastic, probabilistic, or fuzzy models can be applied) the measures of GA can be evaluated as: the expectation (or the fuzzy expectation) of the number of serviced objects by a given time point; the probability (its statistical estimation) of successful service for the given objects.

Similar measures can be proposed for ITA estimations, for example the expectation of the number of objects in a given macro-state, the probability of technological operations fulfillment.

The problem of SF CS GA and ITA evaluation and analysis can be solved on the basis of structure dynamics control models (the model *M* and its components M_o, M_k, M_n, M_e, M_g, M_ν, M_c, M_p.

These models have a form of nonstationary finite-dimensional differential dynamic systems (NFDDS) with reconfigurable structures. So the problem of GA and ITA evaluation can be regarded as a problem of NFDDS controllability analysis. The latter problem, in its turn, can be solved by the NFDDS attainability set $D(t, T_0, \boldsymbol{x}(T_0))$, construction. If the attainability set is obtained, the solvability of the previously stated boundary problems for structure-dynamics control (SDC) can be checked in accordance with the sets of initial X_0 and final X_f states $(\boldsymbol{x}(T_0) \in X_0, \boldsymbol{x}(T_f) \in X_f)$, with the considered period of time, with time-spatial, technical, and technological constraints.

Moreover, the problems of SF SDC stated in the section III can be formulated as follows:

$$J'_{\text{об}}(\mathbf{x}(\cdot)) \to \min_{\mathbf{x}(\cdot) \in D(t_f, t_0, \mathbf{x}(t_0))}, \qquad (13)$$

where $D(t_f, t_0, \mathbf{x}(t_0))$ is the attainability set of the dynamic system (model) M ; $J'_{\mathrm{G}}(\mathbf{x}(\cdot))$ — is the initial functional (12) transformed to the form of Mayer's functional. It is important that the alteration of objective functional does not imply the recalculation of the attainability set $D(t_f, t_0, \mathbf{x}(t_0))$. If the dimensionality of SF CS SDC problem is high, then the construction of the attainability sets becomes a rather complicated problem. Therefore, the approximations of $D(t_f, t_0, \mathbf{x}(t_0))$ ought to be used [Chernousko, Zak, 1985], [Moiseev, 1974], [Sokolov, 1999].

Now we introduce the algorithm of $D(t_f, t_0, \mathbf{x}(t_0))$ construction. The boundary points of the set $D(t_f, t_0, \mathbf{x}(t_0))$ are obtained as the solutions of the optimal control problems [Moiseev, 1974]:

$$J''_{\mathrm{G}}(\mathbf{x}(\cdot)) = \mathbf{c}^{\mathrm{T}}\mathbf{x}(t_f) \to \min_{\mathbf{u}\in Q_p(\mathbf{x})}, \tag{14}$$

where $\mathbf{c}$ is a vector such that $|\mathbf{c}| = 1$. For a given vector $\mathbf{c}$ we obtain the optimal control $\mathbf{u}^*(t)$, the appropriate state vector $\mathbf{x}^*(t_f)$ that is equal to some boundary point of $D(t_f, t_0, \mathbf{x}(t_0))$, and the hyperplane $\mathbf{c}^{\mathrm{T}}\mathbf{x}^*(t_f)$ $\mathbf{c}^{\mathrm{T}}\mathbf{x}^*(t_f)$ to $D(t_f, t_0, \mathbf{x}(t_0))$ at the point $\mathbf{x}^*(t_f)$.

Let $\overline{\overline{\Delta}}$ be the number of different vectors $\mathbf{c}_{\overline{\beta}}$, $\overline{\beta} = 1,\ldots,\overline{\overline{\Delta}}$, then the external approximation $D^{+}(t_f, t_0, \mathbf{x}(t_0))$ of the set $D(t_f, t_0, \mathbf{x}(t_0))$ is a polyhedron whose faces lie on the corresponding hyperplanes, the internal approximation $D^{-}(t_f, t_0, \mathbf{x}(t_0))$ of $D(t_f, t_0, \mathbf{x}(t_0))$ is a polyhedron whose vertices are the points $\mathbf{x}^*_{\beta}(t_f)$, i.e., $D^{-}(t_f, t_0, \mathbf{x}(t_0)) = Co(\mathbf{x}_1(t_f),\ldots,\mathbf{x}_{\overline{\overline{\Delta}}}(t_f))$. The bigger $\overline{\overline{\Delta}}$, the better approximation of the attainability set $D(t_f, t_0, \mathbf{x}(t_0))$ can be obtained. It can be proved [Sokolov, 1999] that the value $\overline{\overline{\Delta}}$ is defined by the total number of possible interruptions for CTS interaction operations over a given time period (t_0, t). To obtain D^{+}, D^{-} Krylov and Chernousko's method was used [Moiseev, 1974]. Instead of the vector $\mathbf{c}$ the vector $\boldsymbol{\psi}(t_0)$ of conjugate variables is to be varied.

Conclusion

The constructed general model has a form of linear (bilinear as regards the model M_k) nonstationary finite-dimensional differential dynamic system with reconfigurable structure. The solutions obtained in the presented multiple-model complex are coordinated by the control-inputs vector $u^{(o)}(t)$ of the model M_o. This vector determines the sequence of interaction operations and fixes SF resources allocation. The applied procedure of solution adjustment was called in [Mesarovic, Takahara, 1975], [Moiseev, 1974], [Tsurlov, 1989] resource coordination.

The model complex M evolves and generalizes the dynamic models of scheduling theory [Moiseev, 1974], [Wanguer, 1969]. The main distinctive feature (as apposed to [Moiseev, 1974]) of the complex is that nonlinear technological constraints are actualized in the convex domain of allowable control inputs rather than in differential equations. Therefore, Lagrangian coefficients keeping the information about technical and technological constraints can be defined explicitly using the local-sections method [Pontriagin, etc., 1961]. In [Moiseev, 1974] more complicated iterative procedure was suggested to obtain Lagrangian coefficients. Furthermore instead of relay constraints $u_v \in \{0,1\}$, $v \in \{1,\ldots,n\}$ (here n is the dimensionality of the control vector u) less strict ones $u_v \in [0,1]$ can be considered. This substitution lets use fundamental scientific results of the modern control theory [Athaus, Falb, 1966], [Bryson, Yo-Chi, 1969], [Fleming, Richel, 1975], [Moiseev, 1974], [Pontriagin, etc., 1961] in various SF control problems (including scheduling theory problems).

Computational investigation [Sokolov, 1999] showed that the use of the SF dynamic models entails considerable dimensionality decrease for control problems to be solved in a real-time operation mode. Recurrence description of models allows parallel computations accelerating problem solving.

The proposed original description of SF control processes establishes dependence relations between control technology and the goals of SF CS. For example, the methods of optimal-control theory applied to the models M_o, M_e, M_n help to estimate the degree of interdependency between quality of spacecraft operating according to a specified purpose and such technological aspects of space-system management as SF resource allocation, trajectory measurement schemes, and information-flow routing methods. Consequently, the optimal programs for resource allocation, for flow routing, and trajectory measurements can be obtained.

Various combinations and interactions of particular control models forming the general model M is the basis for detailed multicriteria analysis of the factors influencing upon the objective results of SF operating.

This research described in this paper is partially supported by grants of Russian Foundation for Basic Research (RFBR) (11-07-00302-a, 10-07-00311-a, 11-08-01016-a, 11-08-00767-a, 12-07-00302). Saint-Petersburg Scientific Center RAS (SPIIRAS project 2012), and the Program of Fundamental Investigation of the Department of Nanotechnologies and Information Technologies RAS (project № 2.11), Program ESTLATRUS project 2.1/ELRI-184/2011/14.

Bibliography

[Athaus, Falb, 1966] Athaus, M., Falb P.L. Optimal control: An Introduction to the Theory and Its Applications. McGrow-Hill Book Company. New York, San Francisco, Sidney, 1966.

[Bryson, Yo-Chi, 1969] Bryson, A.E., and Yo-Chi Ho. Applied optimal control: Optimization, Estimation and Control. Waltham Massachusetts, Toronto, London, 1969.

[Casti, 1979] Casti, J.L.. Connectivity, Complexity and Catastrophe in Large-Scale Systems. Wiley-Interscience, New York and London, 1979.

[Chernousko, Zak, 1985] Chernousko, F.L., Zak, V.L. On Differential Games of Enasion from Many Pursuers // J. Optimiz. Theory and Appl. 1985. Vol.46, N 4. P.461-470.

[Doganovskii, Oseranii, 1990] Doganovskii S.A. Oseranii N.A. Automatic Restructurable Control Systems // Measurement, Cheek, Automation, 1990, N 4 (90). P.62-80 (in Russian).

[Fleming, Richel, 1975] Fleming, W.H., Richel R.W. Deterministic and stochastic optimal control. Springer-verlag, Berlon, New York, 1975.

[Gupta, Sinka, 1996] Intelligent Control Systems: Theory and Applications / Eds. M.M. Gupta, N.K. Sinka. N.Y.: IEEE Press, 1996.

[Klir, 1985] Klir, G.J. Architecture of Systems Problem Solving. Plenum Press, New York. 1985.

[Mesarovic, Takahara, 1975] Mesarovic M.D. and Y. Takahara. General Systems Theory: Mathematical Foundations. Academic Press, New York, Can Francisco, London, 1975.

[Moiseev, 1974] Moiseev, N.N. Element of the Optimal Systems Theory. – Moscow: Nauka, 1974 (in Russian).

[Napolitano, Swaim, 1989] Napolitano M.R., Swaim R.L. A New Technique for Aircraft Flight Control Reconfiguration // Proc. AIAA Guidance, Navigation and Control Conf., 1989. Pt 1, p.1-9.

[Ohtilev, Sokolov, Yusupov, 2006] Ohtilev, M., Sokolov, B., Yusupov, R. Intelligent technologies of monitoring and control. Moscow, Nauka, 2006. 410 p. (in Russian).

[Oreu, Zeigler, Elzas, 1984] Simulation and Model-Based Methodologies: An Integrate View/Ed. by T.I. Oreu, B.P. Zeigler, M.S. Elzas: NATO Series (F), Vol.10 Springer-Verlag Berlin Heidelberg, New York, 1984.

[Pontriagin, etc., 1961] Pontriagin, L.S., etc. Theory of Optimal Process. – Moscow: Fizmatgiz, 1961 (in Russian).

[Roy, 1996] Roy, B. Multicriteria Methodology for Decision Aiding. Kluwer Academic Pulisher, Dordrecht, 1996.

[Russell, Norvig, 1995] Russell, S.J., Norvig, P. Artificial Intelligence: A Modern Approach. Prentice-Hall, Inc., A Simon & Schuster Company, Upper Saddle River, New Jersey, 1995.

[Shannon, 1975] Shannon, R.E. Systems Simulation. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

[Siliak, 1990] Siliak, D.D. Decentralized Control of Complex Systems, Academic Press, New York, 1990.

[Singh, Titli, 1978] Singh, M., and A. Titli. Systems: Decomposition, Optimization and Control, Pergamon Press, Oxford, 1978.

[Sokolov, 1999] Sokolov, B.V., Complex Operations Scheduling and Structure Control in Automation Control Systems of Active Moving Objects. RF. Department of Defense, 1999 (in Russian).

[Sokolov, Kalinin, 1985] Sokolov, B.V., Kalinin V.N., A Dynamic Model and an Optimal Scheduling Algorithm for Activities with Baus of Interrupts // Automation and Telechanics, 1985, N 1, pp.106-114 (in Russian).

[Sokolov, Kalinin, 1995] Sokolov, B.V., Kalinin V.N., Multi-model Approach to the Description of the Air-Space Facilities Control Process // Control Theory and process, 1995, N 1, pp.149-156 (in Russian).

[Tsurlov, 1989] Tsurlov, V.I., Dynamic Problems of large Dimension. – Moscow: Nauka, 1989 (in Russian).

[Tsypkin, 1971] Tsypkin, Y.Z. Adaptation and Learning in Automatic System. Academic Press. New York, 1971.

[Van der Velde, 1984] Van der Velde W.E. Control System Reconfiguration // Proc. American Control Conf., 1984, v.3. P.1741-1745.

[Vasil'ev, 2001] Vasil'ev, S.N., From Classical Control Problems to Intelligent Control // Control theory and process, 2001. N 1. P.5-22. N 2. P.5-21.

[Wanguer, 1969] Wanguer, H.M. Principles of Operations Research. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1969.

[Yusupov, Timofeev, 1994] Yusupov R., Timofeev A. Intellectual systems of automated control // Engineering Cybernetics, 1994. N 5.Pp.211-224 (in Russian).

[Yusupov, Rozenwasser, 1999] Yusupov R., Rozenwasser E. Sensitivity of Automatic Control Systems. CRS. Press, London, New York, 1999.

[Zvirkun, Akinfiev, Filippov, 1985] Zvirkun, A.D., and Akinfiev V.K., Filippov V.A., Simulation Modeling in the Problems of Complex Systems Structure Synthesis. – Moscow: Nauka, 1985 (in Russian).

[Zvirkun, Akinfiev, 1993] Zvirkun, A.D., and V.K. Akinfiev, Structure of the multi-level systems (synthesis and development). – Moscow: Nauka, 1993 (in Russian).

[Zvirkun, 1982] Zvirkun, A.D., Basis of Complex Systems Structure Synthesis. – Moscow: Nauka, 1982 (in Russian).

Authors' Information

Boris Sokolov – Doctor of Sciences (Tech), Prof., Honored scientist of Russian Federation; Deputy-Director for Research of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), 14th line, 39, St.Petersburg, 199178; e-mail: sokol@iias.spb.su.

Major Fields of Scientific Research: Development of research fundamentals for the control theory by structural dynamics of complex organizational-technical systems.

Rafael Yusupov – Doctor of Sciences (Tech), RAS Correspondent member, Honored scientist of Russian Federation; Director of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), 14th line, 39, St.Petersburg, 199178; e-mail: yusupov@iias.spb.su.

Major Fields of Scientific Research: Control theory, informatics, theoretic basics of informatization and information society, information security

Vyacheslav Zelentsov – Doctor of Sciences (Tech), Prof.; Leading Researcher of St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), 14th line, 39, St.Petersburg, 199178; e-mail: v.a.zelentsov@gmail.com .

Major Fields of Scientific Research: Decision-making in complex systems.

MODEL OF LEXICOGRAPHICAL DATABASE: STRUCTURE, BASIC FUNCTIONALITY, IMPLEMENTATION

Olga Nevzorova, Farid Salimov

Abstract: *In the paper we describe the model of lexicographical database and Russian-Tatar lexicographical database data model.*

Keywords: *lexicographical database, linguistic resource, grammatical model.*

ACM Classification Keywords: *H.3 INFORMATION STORAGE AND RETRIEVAL H.3.1 Content Analysis and Indexing - Linguistic processing*

Introduction

Lexicography is the branch of linguistics that deals with the theory and practice of compiling the dictionaries. Theoretical lexicography develops typologies of dictionaries. The dictionaries may differ in form, content and by applied purposes. We can distinguish the following basic functions of dictionaries:

- registration of objective data about the world (encyclopedic dictionaries and defining dictionaries);
- structuring of language content (thesauruses);
- normalization of language as a mean of facilitating communication (standard and terminological dictionaries);
- lexical systematization for language learning (training dictionaries);
- translation from one language to another (bilingual and multilingual dictionaries);
- auxiliary operations.

The development of the dictionary is performed in accordance with a specific concept, which is reflected in the structure or vocabulary or explicitly explained in the introductory article. The use of different parameters in specific lexicographical dictionary is determined by the language specifics, lexicographical traditions, type and purpose of the dictionary and also viewpoints of lexicographers.

The development of new information technologies is contributed to the formation of computational lexicography. The main tasks of computational lexicography are linked with theory and practice of electronic dictionaries. Electronic dictionaries are software products that store dictionary information in a structured way. The development of electronic dictionaries is being performed by appropriate software tools and is being supported by specialized models and methods of text processing. We can distinguish the following types of electronic dictionaries:

complete electronic equivalents of printed dictionaries;

computational lexicographical systems and lexicographical databases based on appropriate lexicographical models.

The electronic version of dictionaries creates new opportunities for the presentation of lexical material including the possibility of partial presentation according to different criteria (different projections of the dictionary). The electronic version allows to apply different linguistic technologies (morphological and syntactic analysis, full-text search, speech recognition and speech synthesis, etc.) to the content of articles. Lexical entry can contain a variety of grammatical and semantic features, phonetic information, lists of translations, list associated with the key entries (collocations, standard phrases, etc.), text comments, synonymic and antonymic rows . Examples of successful commercial solutions in the field of computational lexicography are electronic dictionaries of “Lingvo” family of ABBYY company.

Lexicographical database and lexicographical systems

Technology development of modern information systems are closely linked with the technology of database development. The theory of database is a fully established discipline, based on formal approaches, methods and technologies of creating and maintaining the structured data, as well as finding information in large distributed datasets. To create and maintain a database (updates, access to records on request and issue records to the user) a database management system (DBMS) is used. Description of the database at the conceptual level is a generalized view of the data in terms of subject area (the application developer, user, or an external information system). The data model defines the rules of data structuring in the database. The data model includes a set of principles and methods for describing data and methods for manipulating data. According to the type used by the data model there are three common classes of databases: hierarchical, network and relational. By functionality, one can be identified as operational and reference information databases. The last one includes catalogs of electronic libraries, electronic dictionaries, statistical databases, etc. These systems are used to support core business processes and does not require changes to existing records. Operational databases are designed for more control of various technological processes. In this case, data are not only retrieved from the database, but data are changed (added), including the result of this usage. In the field of possible applications one can distinguish between universal and specialized (or problem-oriented) systems.

Lexicographical database (LDB) contains data on different levels of linguistic units (from morphemes to the text) and a variety of information about these units. One can specify the following applications for the LDB:

- supporting of operation of various automated systems associated with the processing of text and speech (information systems, expert systems, training systems, speech analysis, machine translation, etc.);
- computerized lexicographical tools for development of dictionaries of different types (training dictionaries, translation dictionaries etc.)
- automation of activities of researchers: linguists, language teachers and other professionals.

Structuring information in a database is a tool to automate the process of search of information in large data. Searching of information in the context of a database is considered as a sequence of operations to find objects with certain properties. One of the major problems is the problem of constructing an effective search in the databases. In [Ryzhov, 2004] the models of strict and fuzzy databases are introduced and the models of strict and fuzzy queries to the databases are developed. According to this classification, a strict database deals with a set of records whose values uniquely are interpreted by users. The strict request is regarded as a logical expression with the logical terms that are expressed by the usual means of set theory To define request, it is possible to enumerate the values of attributes or objects or to indicate the change of parameters and attributes and to associate data of pair "attribute-value" with logical connectives.

Fuzzy query, in contrast to the strict query, may contain logical terms with fuzzy values. For example, the values of the attribute "size" may be "small", "very little", "big", etc. In this case, the boundary between the values of a fuzzy query is not defined, and each user has their own idea about these limits. This is the difference between a strict and fuzzy databases: the attributes of the fuzzy ones may be fuzzy values. Searching for a strict request in strict databases is well designed. To solve the problem of effective search one must be able to construct the predicates using first-order predicate logic, which take true values on the set of data interested. Execution time optimization should be carried out under constructing the query. Query is focused on searching the set of data that satisfy certain conditions. It is important to choose a formalized description language of the target set under constructing such queries. For this purpose, different tasks use different means: the language of predicate logic, the language of regular expressions, etc. For example, conjunctive normal form (CNF) as the basic formula for writing a query is used in the Russian National Corpus [Ruscorpora]. The user selects the required characteristics from a set of attributes on understandable language under constructing the query. The system based on the selected values, builds a formula of query. In particular, one can search a set of lexical items for a given set of grammatical or semantic attributes. The structure of the base relations has important role under performance the query. When designing a database it is necessary to consider the structure and types of possible queries. For fuzzy queries the problem of describing the boundaries of the fuzzy variables (for example, what value of size can be interpreted as a "big") arises. In [Ryzhov, 2004] the approach to the formation of such requests involving the theory of fuzzy sets is discussed.

In [Shirokov, 1998] the concept of lexicographical system is described. The lexicographical system is regarded as an information environment in which lexicographical models are implemented. Special cases (or implementations) of lexicographical systems are the systems of description language (both computational and traditional systems). The dictionary as an abstract lexicographical system must have a structure that includes at least two essential parts - the left part (registry) and right part (interpretation part). The vocabulary has interpretation unlike word list. But the dictionary has a deeper structure, which appears in the structure of the left and right parts of the dictionary as a whole and its entries, as well as in the structure of links between dictionary entries or links between different parts of dictionary. Thus, the dictionary is a special kind of text, which is a description of the vocabulary in a systematic and structured way. Each dictionary entry begins with a dictionary headword. The set of headwords forms the vocabulary (or the left part of the dictionary). The choice of vocabulary depends on the purpose of a dictionary.

Vocabulary may consist of the following language units:

- phonemes;
- morphemes (prefixes, roots, suffixes, etc.) - for a dictionary of morphemes, grammatical dictionaries, dictionaries of word formation;
- lexemes - for the majority of dictionaries (monolingual, spelling, etc.);
- word forms - for grammatical dictionaries, dictionaries of rhymes, etc.;
- collocations - for phraseological dictionaries, dictionaries of idioms, dictionaries of clichés.

The right part of the vocabulary is used to explain the headword. Right part zones are designed for each dictionary individually. The right part may contain a list of synonyms of the word, the translation of words (for dictionaries of foreign words), the interpretations of the concept, which describes the given word, and different applications (graphs, charts, figures, etc.). The set of all entries forms corpus of the vocabulary.

The structure of the dictionary entry can be quite complicated and has a large number of structural elements. Since the structure of the article defines a system of basic relations for the database, it is important formal description of such a structure. A formal approach to the description of the various linguistic objects helps to build the algorithms for the selection of objects based on the verification of certain conditions and predicates, which are defined in the model description.

Models of lexicographical systems and lexicographical databases can be considered as information systems and models for describing data and methods for manipulating data are very important. In this case the main lexicographical work associated with the isolation and descriptions of lexical units are performed by professional linguists although their professional activities may be supported by a set of specialized software tools. However, it should be clearly understood that the selected material requires special linguistic formalization for the effective integration of database technology. The next section will discuss a model describing the data and methods for performing searches on the developed data model, made in the design of the Russian-Tatar lexicographical database.

Russian-Tatar lexicographical database data model

At the moment, Russian-Tatar lexicographical database is being developed at Research Institute for Applied Semiotics at Tatarstan Academy of Science. Main aim of this project is developing baseline resource for linguistics software used for knowledge intensive science projects, such as Tatar language corpus and parallel corpus, machine translation and others. Developed solutions should have effective functionality, which supports storing and searching of multiparameter linguistic definitions. An important part of this is to implement extensions for inclusion of new languages (firstly Turkic languages with similar structures).

This linguistic resource can be characterized as having very fine levels of specification of linguistic markup and models internal relationships inside of and between Tatar and Russian languages lexical systems. Database consists of interlinked components, corresponding to Russian and Tatar languages. Each of these components have independent internal structure, caused by specifics of the language in question and linked on lexical equivalency level. Each language component is represented by grammar and semantic models.

Process of designing Russian-Tatar lexicographical database was geared towards solving several major problems:

- Building data definition schema based upon linguistic models;
- Development of coding scheme for encoding linguistic data to avoid duplication of source data;
- Development of effective search mechanisms for accessing data.

LDB structure is defined by theoretical linguistic models that are being used, also by formalization level of such models and by an ability of these models to be expressed in the database level logic.

Information retrieval tasks are comprised of several important problems, such as searching for grammatical attributes of a given word form (direct search problem), and reverse problem of searching for a set of lexical tokens, whose grammatical attributes include those that supplied by user. Solving direct search problem is equivalent to solving morphological analysis problem over set of word forms, which are defined by dictionary.

Special properties of Tatar language defined a series of decisions, made in process of designing representation for its grammatical model (T-component). One of Tatar languages defining characteristics is separation of any word form into root and affix parts, with affixational part defining word form's morphologic attributes. All lexical tokens of Tatar language can be divided into four morphological types, according to which affixational morphemes can linked with base morphemes of specific type. Same morphological attribute can be characterized using different allomorphs, which can joined into morphological category classes according to their morphological and morphonological types. Because of that information about word form was divided into two parts: dictionary of base morpehemes and dictionary of inflexional suffixes. Dictionary of base morphemes stores information about lemmas and morphological and morphonological types. Dictionary of inflexional suffixes contains possible chains of affixes which are linked with dictionary of base morphemes based on types mentioned earlier. In theory, those chains can have infinite length because of duplicate morphemes. Research of the statistical data has shown that in Tatar language word forms with more than 5 affixes make up less than 1% of word forms encountered in texts. Because of that property, length of chains in dictionary of inflexional suffixes is capped at 5 allomorphs. Also linked with this dictionary is table of morphological categories, which contains information about rules for building affixational chains, in addition to decoded values of corresponding morphological attributes.

The figure 1 depicts part of database structure, linked to T-component. T_Base table contains information about lemmas, T_Okon describes possible affixal chains, which in turn are divided into morphological categories (M_Posled table). Attributes WG and FORMA describe morphological and morphonological types of lemma accordingly. Relationship between T_Base and T_Okon, defined by WG and FORMA attributes, is M:M.

In the case of database being constructed according to such principles, implementation of direct search of morphological attributes is rather trivial: you just need to split word form into root and inflexional suffix and retrieve morphological attributes based on suffix alone.

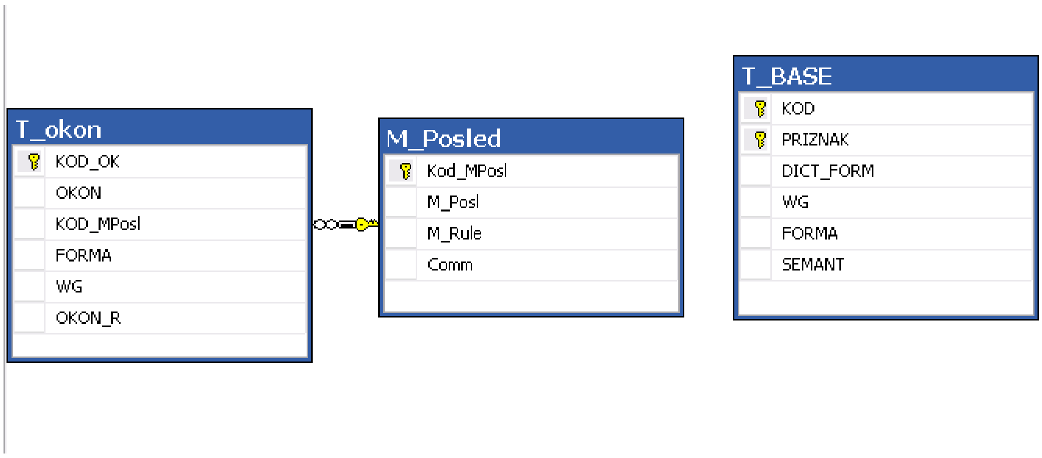


Fig. 1. The fragment of lexicographical database structure

Complications arise if you try to implement reverse search algorithm along same lines as a direct one. Full list if of grammatical attributes can be quite extensive, it's elements interlinked with others according to specific hierarchical relationships, each lexical token with its own set of morphological attributes. Storing full set of attributes in database using standard RDBMS practices (each attribute has its own field) can be quite ineffective, mainly because in majority of cases values of attributes won't be equal and so resulting database will be sparse and riddled with empty values. In such cases it becomes necessary to use specially designed encoding methods adjusted to search techniques used. Simplest solution is to use encode information using characteristic vectors, consisting of 0 and 1 depending on membership of specific attribute within set that is being encoded. Set of attribute values is ordered according to specific condition and each affixal chain has its own characteristic vector. Retrieving set of word forms with specified set of attributes is implemented as matching resulting vector $\overline{b}$ with rows in table of morphological categories. If condition $\overline{b} \le \overline{a}$ is fulfilled, whereas $\overline{a}$ is characteristic vector for a sought morphological category, then search is stopped. Considering that for binary vectors $\overline{b} \le \overline{a}$ is equivalent to $\overline{b} = \overline{a} \,\&\, \overline{b}$, whole process can be implemented using only bitwise operations. Described approach has added benefit of accommodating further expansion of grammatical parameters set: arbitrary number of additional bits can be reserved in advance and adding data about new parameter can be as simple as switching 0 to 1 in appropriate position. Another approach is to split characteristic vector into parts $\overline{a_i}$ and $\overline{b_i}$, $i = \overline{1,k}$ with modified predicate $\&_{i=1}^{k} ((\overline{a_i} \,\&\, \overline{b_i}) = \overline{b_i})$. Main disadvantage of using bit vectors to encode attribute data is that each attributes position is fixed in accordance with enumeration order. It matters because of "recursive" inflectional suffixes in Tatar language which require keeping track of grammatical attribute's repetition factor in word form. For example, consider attributive affix - *Dagi* and possessive affix – *Nyky.* These affixes can be joined multiple times with root part of the word, potentially creating "cyclic" word form of theoretically infinite length (in Tatar *urman+nar+ym+dagi+lar+ym+dagi* … «…those in those that in my forests.. »). In such cases, it is necessary to reserve sufficient space in characteristic vector for repetition of attributes in affixational chain.

Another approach for encoding morphological data treats morphological attributes as formulas of propositional calculus. As mentioned earlier, sets of attributes forms a set of ordered structures, conjunctive normal forms can be used as an expressions. Each conjunction can be mapped to a definite grammatical characteristic and, being comprised of disjunctions, evaluates a set of values of respective attribute. Using this approach to encoding reduces searching to a comparison between two different formulas. Each formula A implements some function f_A, defined on some subset of grammatical attributes. Let's define that formula A covers formula B $(A \supseteq B)$ if for any subset of grammatical attributes over which $f_B = 1$ also true that $f_A = 1$. If we use conjunctive normal form, it is easy to introduce conditions for verifying coverage of a function by another. Assume that *B* is a formula, created from a targeted request from a user. Executing a search against this formula will result in set of word forms, whose linked formulas cover *B.*

Described approach is flawed because formula attribute breaks 1NF atomicity requirement. This requires checking coverage relationship using in-database defined functions. Checking process can be sped up by introducing linear ordering on conjunctions and disjunctions using weight function (which uses frequency distribution of an attribute in all formulas) over set of grammatical attributes and terminating comparison process if coverage breach is discovered.

At the moment second approach is implemented in the database. Encoded CNF of attributes is stored in *M_Posl* field of *M_Posled* table. Linguistic database is implemented using MS SQL Server 2005. Tests on database consisting of 4000 lexical tokens have shown that query with formulas that include about 50 grammatical attributes takes 3ms to complete on 2.8GHz processor with 512MB RAM and internal caching enabled.

In conclusion we'd like to note that proposed approach to encoding grammatical components is language-independent and was used in processing Russian component regardless of completely different database schema. This approach can also be used for searching word forms with specified set of grammatical attributes.

Conclusion

Development of database for lexicographical models can be difficult task, requiring both comprehensive practical and theoretical knowledge of database systems with knowledge about lexicographical models and specific features of language structures. As a result, trade-offs in design process are inevitable, such as breaking atomicity requirement and introducing additional value processing procedures, which significantly complicates development process.

Design of structures and functional features of LSD is based on the specific problems of lexicographical data processing. Models database is focused primarily on the direct retrieval of data (find the characteristics of a given object), while at the same time, the task of goal-driven search (find all of the objects with given features) is very important for linguistic research. Combining in one data model effective solutions to the forward and reverse search is a difficult task.

The article is suggested quite efficient solutions of the above problems for the project Russian-Tatar lexicographical lexicographical database. In the future, we will plan to develop given LSD in the expansion of the semantic descriptions of vocabulary, as well as expanding to other Turkic languages, based on the description of grammatical formalisms proposed models.

Acknowledgements

The work was supported by the Russian Scientific Foundation for the Humanities, project No 11-14-16024a / B.

Bibliography

[Ryzhov, 2004] Ryzhov A.P. Modeli poiska informacii v nechetkoj srede. M. MGU, 2004. In Russian.

[Ruscorpora] Ruscorpora [Electronic resource] http:\\www.ruscorpora.ru/

[Shirokov, 1998] Shirokov V.A. Informacijna teorija leksikografichnih sistem. Kiïv, Dovira, 1998. – 331p. In Ukrainian.

Authors' Information

***Olga Nevzorova** – Vice-director of Research Institute of Applied Semiotics of Tatarstan Academy of Sciences, Kazan Federal University. P.O. Box: 420111, Bauman str., Kazan, Russia; e-mail: onevzoro@gmail.com*

Major Fields of Scientific Research: Natural language processing, Artificial intelligence

***Farid Salimov** – Senior Scientific Researcher of Research Institute of Applied Semiotics of Tatarstan Academy of Sciences, Kazan Federal University. P.O. Box: 420111, Bauman str., Kazan, Russia; e-mail: Farid.Salimov@ksu.ru*

Major Fields of Scientific Research: Software technologies, Multi-dimensional information systems

PROBABILISTIC ESTIMATION OF TRUST MODEL AND THREAT RESISTANCE ANALYSIS IN SERVICE-ORIENTED SYSTEMS

Nataliia Kussul, Olga Kussul, Sergii Skakun

***Abstract**: Trust and reputation models play an important role in enabling trusted computations over large-scale distributed Grids. Many models have been recently proposed and implemented within trust management systems. Nevertheless, the existing approaches usually assess performance of models in terms of resource management while less attention is paid to the analysis of security threat scenarios for such models. In this paper, we asses the most important and critical security threats for a utility-based reputation model in Grids. The existing model is extended to address these threat scenarios. Also we propose the probabilistic estimation of trust model. With simulations that were run using data collected from the EGEE Grid-Observatory project, we analyze efficiency of the utility-based reputation model against these threats.*

***Keywords**: trust; reputation model; Grid computing; utility; security threats*

***ACM Classification Keywords**: H.1.1 [Models and Principles] Systems and Information Theory; I.4.8 [Image Processing and Computer Vision] Scene Analysis - Sensor Fusion*

Introduction

Grid represents a distributed environment that integrates heterogeneous computing and storage resources administrated by multiple organizations. One of the main concepts in Grid is a virtual organization (VO) — a set of individuals and/or institutions defined by coordinated resource sharing rules for reaching common goals (Foster et al., 2001). VOs are formed dynamically, exist for some time and then resolve.

Trust and reputation models play an important role in enabling trusted computations over large-scale distributed Grids. Two types of trust management systems (TMSs) can be discriminated (Chakrabarti, 2007): policy-based and reputation-based. In policy-based systems, entities in a VO establish trust relationships based on certain predefined policies. In reputation-based systems, certain mechanisms exist in order evaluate the trust which is the function of reputation. Reputation can be viewed as an assumption about the expected quality or reliability of a resource based on existing information or observations about his behaviour in the past (Abdul-Rahman and Hailes, 2000).

Many trust and reputation models have been recently proposed for distributed systems and for Grids, in particular (Arenas et al., 2008; Azzedin and Maheswaran, 2002; Eymann et al., 2008; Gomez Marmol and Martınez Perez, 2008; Josang et al., 2007; Kamvar et al., 2003; Kerschbaum et al., 2006; Liang and Shi, 2010; Papaioannou and Stamoulis, 2008; Silaghi et al., 2007; Song et al., 2005; Srivatsa and Liu, 2006; von Laszewski et al., 2005; Wu and Sun, 2010). Nevertheless, the existing approaches usually assess performance of models in terms of resource management while very few of them focus on the analysis of security threat scenarios for such models.

Gomez Marmol and Martınez Perez (2009) described security threats scenarios in trust and reputation models for distributed systems and proposed possible solutions to tackle them. The study also shows how some of the most representative models (mostly for P2P systems) deal with those threats. von Laszewski et al. (2005) extended an EigenTrust model (Kamvar et al., 2003) to be used in Grids (GridEigenTrust). The obtained reputation value is integrated into a QoS management system providing a way to re-evaluate resource selection and service level agreement (SLA) mechanisms. Eymann et al. (2008) investigated economical issues in Grids along with information asymmetry. These issues are taken into consideration while proposing a reputation-based framework for enabling Grid markets and allowing grid service broker to deal effectively with hidden information. Srivatsa and Liu (2006) identified vulnerabilities that are crucial to decentralized reputation management and developed a safeguard framework for providing a highly dependable and efficient reputation system, called TrustGuard. The conducted experiments showed that the TrustGuard framework is effective in countering malicious nodes regarding oscillating behaviour, flooding malevolent feedbacks with fake transactions, and dishonest feedbacks.

In this paper, we asses the most important and critical security threats for a utility-based reputation model in Grids that was proposed by Silaghi et al. (2007) and Arenas et al. (2008). We will use security threat scenarios for trust and reputation models presented by Gomez Marmol and Martınez Perez (2009) as a reference in our study. These scenarios include: individual malicious peers, malicious collectives, malicious collectives with camouflage, malicious spies, Sybil attack, man in the middle attack, driving down the reputation of a reliable peer, partially malicious collectives, and malicious pre-trusted peers. The model is further extended to address these threat scenarios. With simulations that were run using data collected from the EGEE Grid-Observatory project (Germain-Renaud et al., 2011), we will analyze efficiency of the utility-based reputation model against these threats.

Utility-based reputation model for VOs in Grids

In this paper we extend the existing utility-based reputation model (Arenas et al., 2008) by incorporating a statistical model of user behaviour (SMUB) that was previously developed for computer networks and distributed systems (Kussul and Skakun, 2004; Shelestov et al. 2008, 2007; Skakun et al., 2005) and several new components to address security threat scenarios. The proposed extensions to the reputation model include:

- assigning initial reputation to a new entity in VO: when organization provides a new resource to be integrated in a VO there are no records from the monitoring system to infer reputation value for this specific resource. One possible way of assigning initial reputation to a new resource is to use a methodology of active experiment. There can be several benchmark tasks in the system to estimate the utility function and to provide initial reputation of the resource.

- alliance between consumer and resource: since reputation of resource is based on measure of satisfaction of a consumer in relation to this resource we should avoid cheating via collusions among a group of entities (Azzedin and Maheswaran, 2002). For this purpose, it is advisable to include into the model a factor that will reflect alliance between the consumer and resource.

- time decay function: reputation of resource is based on measuring average value of utility function over certain period of time (Azzedin and Maheswaran, 2002; Silaghi et al., 2007). But if a VO exists for a considerable period of time (e.g. for years) reputation of resource may vary considerably. That is, it is unlikely to use, for example, two years data to estimate current resource reputation if more recent records are available. So, we propose to incorporate a time lag function into the model that will provide weights depending on the time of the transaction record between consumer and resource.

- score function: for different types of services offered by resource providers different reputation values will be used (Gomez Marmol and Martınez Perez, 2009). Namely, we will categorize services into categories, and a resource provider will get reputation value according to such a category. In Grid systems, tasks can be categorised by the computational complexity. Successful execution of tasks with a complex workflow and parallel programs, for example, environmental models like numerical weather prediction (Kussul et al., 2009; Hluchy et al., 2010), will provide to a resource provider higher reputation value.

Reputation model for resource providers

For the reputation model we will use the enhancement of the well-known model (Arenas et al., 2008), proposed in (Kussul, Novikov, 2009).

The reputation model is based on the utility function that measures the level of satisfaction of a user in relation to service provider. In order to define utility function an auxiliary function that indicates the SLA accorded between a VO user and a resource provider for a particular resource within a VO is implemented (Arenas et al., 2008):

$$\text{SLA} : \bigcup_l u \times \bigcup_k r_k \times \bigcup_m vo_m \to \text{R} \qquad (1)$$

where R denotes the set of real numbers.

The SLA value represents quality of resource provider as expected by user (Arenas et al., 2008). In order to define utility function based on SLA value we describe the notion of Event:

$$\text{Event} = \text{T} \times \bigcup_l u_l \times \bigcup_k r_k \times \bigcup_m vo_m \times \{\text{QoS name}\} \times \text{R} \qquad (2)$$

where T is a time domain.

Before defining utility function and reputation we will introduce three functions: the first one will characterise possible alliance between consumer and resource in order to avoid cheating (Azzedin and Maheswaran, 2002), the second one will account for a time when utility was estimated (Azzedin and Maheswaran, 2002; Silaghi et al., 2007), and the third one will provide different scores depending on the type of the provided service (Gomez Marmol and Martınez Perez, 2009). These functions provide extensions to the utility function and reputation originally proposed by Arenas et al. (2008).

Function h(u, r) will take a value between 0 and 1 and will show the level of alliance between user u and resource r. If there is no such an alliance between targets, h(u, r) will have a higher value. For example, one possible way of defining h(u, r) is as follows

$$h(u,r) = \begin{cases} 1, \text{ if } f_{vo}(r) \neq g_{vo}(u) \\ \theta, \text{if } f_{vo}(r) = g_{vo}(u) \end{cases}, \quad (3)$$

where θ is a parameter.

Function z(*t, tc*) will show what past records on user-resources interactions should be taken into consideration to estimate reputation of specific resource. Here t is the time, and tc is a parameter. In a simplest form z(t, tc) could be a stepwise function

$$z(t,t_c) = \begin{cases} 1, t \ge t_c \\ 0, t < t_c \end{cases}. \quad (4)$$

Function s(type(r)) will provide different values for different types of services provided by the resource r (function type(r) maps into category of service).

Now, we can define a utility function

utility : *Event* → R,

$$utility(\{t, u, r, vo, QoS, v\}) = \begin{cases} h(u,r)s(type(r)), & \text{if } SLA \text{ met} \\ penalty(\nu, SLA)h(u,r)s(type(r)), & \text{otherwise} \end{cases}, \quad (5)$$

where *SLA* is the agreed *SLA* value between the user and resource provider, *penalty*(*v*, *SLA*) is a penalty function imposed on a resource provider if the agreed SLA is not met.

The form of penalty function depends on the QoS in place. For example, for time metrics which are usually to be minimised a penalty function can be represented by

$$penalty(v, SLA) = \begin{cases} 1, & \text{if } \nu \le SLA \\ \frac{SLA}{v}, & \text{if } \nu > SLA \end{cases}. \quad (6)$$

Let us denote a set of traces that are used to estimate the reputation of resource r in a vo up to the current time t with

Trace|(vo, r, t) = $\{\{t', u', r', vo\_id', QoS', \nu'\} \in Trace : r = r', vo\_id = vo', t' \le t\}$. (7)

Let us denote a set of utility() function values derived from traces Trace|(vo, r, t) with

O(vo, r, t) = { z(t, tc)·utility({t, u, r, vo, QoS, v}) | {t, u, r, vo, QoS, v}$\in$ Trace|(vo, r, t)}. (8)

A reputation is expectation of utility() function (in terms of probability theory)

rep(vo, r, t) = E[utility(O(vo, r, t))] = $\int utility(O_{(vo,r,t)})\, p_{utility}(O_{(vo,r,t)})\, dO_{(vo,r,t)}$. (9)

If we do not want to discriminate values from utility() function by time then we might use z(t, tc) = 1.

In order to approximate expectation we can use a sample mean

$$rep(vo, r, t) = \frac{1}{|O_{(vo,r,t)}|} \sum_{x \in O_{(vo,r,t)}} x, \quad (10)$$

where | · | denotes the cardinality of the set.

The reputation of an organisation o in VO is the aggregation of the reputation of all resources it provides to VO:

$$rep(vo, t) = \frac{1}{|f_{vo}^{-1}(o)|} \sum_{r \in f_{vo}^{-1}(o)} rep(vo, r, t). \quad (11)$$

The reputation of a resource in all VOs can be estimated as follows

$$rep(r, t) = \frac{1}{|VO|_r|} \sum_{vo \in VO|_r} rep(vo, r, t). \quad (12)$$

Probabilistic reputation based trust model

Let's describe our model in terms of the theory of probability to enable the theoretical analysis of its properties and limitations, as well as assessing the security of the model against the threat scenarios. If *SLA* is a Service Level Agreement, *v* stands for the actual value of the provided services (obtained after the service has been provided to the user). We will denote as *ξ* the random value that shows the agreed SLA, also we will denote the meaning of *v* as *η*. After that we can define the penalty function *penalty*(*v*, *SLA*) and the corresponding random value *θ* as follows:

$$penalty(v, SLA) = \frac{SLA}{v}, \ \theta = \frac{\xi}{\eta}, \tag{13}$$

We will calculate distribution function of the random value *θ* through the corresponding functions of the variables ξ and *η* (provided that *ξ*, *η* > 0):

$$P\{\theta < z\} = P\left\{\frac{\xi}{\eta} < z\right\} = \iint_{\mathrm{A}} p_\xi(y)p_\eta(x)dxdy = \iint_{0<y<k_x} p_\xi(y)p_\eta(x)dxdy =$$
$$= \int_0^\infty \int_0^{kx} p_\xi(y)p_\eta(x)dxdy = \int_0^\infty p_\eta(x)\int_0^{kx} p_\xi(y)dydx, \tag{14}$$

where $\mathrm{A} = \left\{(x,y): \frac{y}{x} < z, x > 0, y > 0\right\}$.

In case if *SLA* value for the specific service is constant, then we can present (14) as follows:

$$P\{\theta < z\} = P\left\{\frac{SLA}{\eta} < z\right\} = P\left\{\eta > \frac{SLA}{z}\right\} = \int_{\frac{SLA}{z}}^\infty p_\eta(x)dx \tag{15}$$

According to (10) the utility function:

$$u = \begin{cases} 1, \text{ if } \theta \ge 1 \\ \theta, \text{ if } \theta < 1 \end{cases}. \tag{16}$$

In this case, the distribution function of the random value *u* will be defined as follows:

$$\begin{aligned} &P\{u = 1\} = P\{\theta \ge 1\} \\ &P\{u < x\} = P\{\theta < x\}, \ \ where \ 0 \le x < 1 \end{aligned}. \tag{17}$$

Reputation is the mathematical expectation of utility function:

$$rep = \mathrm{M}[u] = \int_0^1 xp_u(x)dx, \tag{18}$$

where $p_u(x)$ - density function of the random value *u*.

Lets calculate this expression.

$$\int_0^1 xp_u(x)dx = 1 \cdot P\{\theta \ge 1\} + \int_0^1 xp_\theta(x)dx = P\{\theta \ge 1\} + \int_0^1 xp_\theta(x)dx =$$
$$= P\{\theta \ge 1\} + \mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}],$$

where $\mathbf{1}_{\theta<1} = \begin{cases} 1, & if \;\; \theta < 1 \\ 0, & otherwise \end{cases}$.

That's why, the reputation of the resource can be estimated as follows:

$$rep = \mathrm{M}[u] = P\{\theta \ge 1\} + \mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}]. \qquad (19)$$

The first summand $P\{\theta \ge 1\}$ shows the probability that the resource will fulfill the SLA, the second summand $\mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}]$ shows the mean value of the penalty function, if SLA will be violated.

Lets look at the following example. Let the *SLA* be a fixed value (in this case the parametric variable) and the random variable *η* is distributed according to Pareto distribution, so

$$P\{\eta < x\} = \begin{cases} 1 - \left(\dfrac{x_m}{x}\right)^{\alpha} & \text{if } x \ge x_m \\ 0 & \text{if } x < x_m \end{cases}. \qquad (20)$$

According to (15) and (20):

$$P\{\theta < z\} = P\left\{\eta > \frac{SLA}{z}\right\} = \begin{cases} \left(\dfrac{z \cdot x_m}{SLA}\right)^{\alpha} & \text{if } z \le \dfrac{SLA}{x_m} \\ 1 & \text{if } z > \dfrac{SLA}{x_m} \end{cases} \text{ and}$$

$$p_{\theta}(z) = \begin{cases} \alpha\left(\dfrac{x_m}{SLA}\right)^{\alpha} z^{\alpha-1} & \text{if } z \le \dfrac{SLA}{x_m} \\ 0 & \text{if } z > \dfrac{SLA}{x_m} \end{cases}.$$

According to this expression:

$$P\{\theta \ge 1\} = P\{\eta \le SLA\} = \begin{cases} 1 - \left(\dfrac{x_m}{SLA}\right)^{\alpha} & for \;\; SLA \ge x_m \\ 0 & for \;\; SLA < x_m \end{cases}.$$

The resource reputation assessment can be divided into the following cases:

1. If $\dfrac{SLA}{x_m} \le 1$, then $P\{u = 1\} = P\{\theta \ge 1\} = 0$. This scenario describes the resource that is always providing a bad service. It means that this resource never meets SLA. Let's assess the reputation of such a resource.

$$rep = P\{\theta \ge 1\} + \mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}] = \mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}] = \int_0^1 x p_{\theta}(x) dx =$$

$$= \int_0^{\frac{SLA}{x_m}} x\alpha\left(\frac{x_m}{SLA}\right)^{\alpha} x^{\alpha-1} dx = \alpha\left(\frac{x_m}{SLA}\right)^{\alpha} \int_0^{\frac{SLA}{x_m}} x^{\alpha} dx = \alpha\left(\frac{x_m}{SLA}\right)^{\alpha} \left.\frac{x^{\alpha+1}}{\alpha+1}\right|_0^{\frac{SLA}{x_m}} =$$

$$= \alpha\left(\frac{x_m}{SLA}\right)^{\alpha} \frac{\left(\dfrac{SLA}{x_m}\right)^{\alpha+1}}{\alpha+1} = \frac{\alpha}{\alpha+1}\frac{SLA}{x_m}.$$

Therefore,

$$rep = \frac{\alpha}{\alpha+1}\frac{SLA}{x_m}. \qquad (21)$$

2. If x_m =0, then $P\{u=1\} = P\{\theta \geq 1\} = 1$. This scenario describes the resource that is always providing a bad service. It means that this resource always meets SLA. The reputation of such a resource is 1, because $rep = P\{\theta \geq 1\} + \mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}] = 1$.

3. If $\frac{SLA}{x_m}$ >1, then $P\{u=1\} = P\{\theta \geq 1\} = 1 - \left(\frac{x_m}{SLA}\right)^{\alpha}$. This scenario describes the resource that is always providing a partially unreliable service. It means that in some situations the agreed SLA is met by the resource and in others violated. Let's assess the reputation of such a resource.

$$rep = P\{\theta \geq 1\} + \mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}] = 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} + \int_0^1 x p_\theta(x) dx =$$

$$= 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} + \int_0^1 x\alpha \left(\frac{x_m}{SLA}\right)^{\alpha} x^{\alpha-1} dx = 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} + \alpha \left(\frac{x_m}{SLA}\right)^{\alpha} \int_0^1 x^{\alpha} dx =$$

$$= 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} + \alpha \left(\frac{x_m}{SLA}\right)^{\alpha} \left.\frac{x^{\alpha+1}}{\alpha+1}\right|_0^1 = 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} + \alpha \left(\frac{x_m}{SLA}\right)^{\alpha} \frac{1}{\alpha+1} =$$

$$= 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} \frac{1}{\alpha+1}.$$

Therefore,

$$rep = 1 - \left(\frac{x_m}{SLA}\right)^{\alpha} \frac{1}{\alpha+1}. \qquad (22)$$

The obtained results are summarized in Table 1.

Table1 — Different service types and reputation, calculated for Pareto distribution of QoS metrics with x_m and α parameters and fixed SLA value.

Resource type	Parameters	$P\{\theta \geq 1\}$	$\mathrm{M}[\theta \cdot \mathbf{1}_{\theta<1}]$	Reputation
Always good service	$\frac{SLA}{x_m} \leq 1$	0	$\frac{\alpha}{\alpha+1}\frac{SLA}{x_m}$	$\frac{\alpha}{\alpha+1}\frac{SLA}{x_m}$
Always bad service	x_m =0	1	0	1
Partially unreliable	$\frac{SLA}{x_m}$ >1	$1 - \left(\frac{x_m}{SLA}\right)^{\alpha}$	$\left(\frac{x_m}{SLA}\right)^{\alpha} \frac{\alpha}{\alpha+1}$	$1 - \left(\frac{x_m}{SLA}\right)^{\alpha} \frac{1}{\alpha+1}$

Let's analyze the obtained reputation values in terms of the parameters value. If $\frac{SLA}{x_m} \to 1+$ (tends to 1 on right), in this case partially unreliable resource always provides bad service: when $\frac{SLA}{x_m} = 1$ the reputation of this two resources equals to $\frac{\alpha}{\alpha+1}$. When $x_m \to 0$ and the SLA value is fixed, then in this case partially unreliable resource always provides good service, because $1 - \left(\frac{x_m}{SLA}\right)^{\alpha} \frac{1}{\alpha+1} \to 1$. We can get the same result if we fix x_m and $SLA \to \infty$: $1 - \left(\frac{x_m}{SLA}\right)^{\alpha} \frac{1}{\alpha+1} \to 1$.

For the resource that always provides bad service: if $x_m \to \infty$ or $SLA \to 0$ reputation $\frac{\alpha}{\alpha+1} \frac{SLA}{x_m} \to 0$.

Analysis of security threat scenarios for utility-based reputation model

Usually reputation models are analysed in terms of performance, for example resource management, while less attention is paid to the analysis of security threat scenarios. In this section we will study different security threats scenarios in the area of trust and reputation management that were proposed by (Gomez Marmol and Martınez Perez, 2009), and analyse how the proposed model responds to these threats. It should be noted that some of these attacks can be handled by existing mechanisms already implemented for Grids.

1 Individual malicious peers

Malicious peers always provide bad services (Gomez Marmol and Martınez Perez, 2009). From Grid perspective, there can be either a resource that always provides unreliable services, or a malicious user that always tries to harm a system. Such an unreliable resource will provide poor services to the users that will result that the agreed SLA would not be always met (for example, $v << SLA$ for time-related QoS metrics), and thus the reputation of this resource will be always low.

2 Malicious collectives

This is a situation when malicious peers that always provide bad service form a malicious collective (Gomez Marmol and Martınez Perez, 2009). In Grids, there could be a user that tries illegally to improve the reputation of a particular resource. If the user and resource belong to the same organization that kind of behaviour will be captured by the alliance function h(u, r). In order to improve the reputation value considerably the user will need to submit a lot of simple jobs. (Here, by simple jobs we mean jobs that would not require much CPU time and will be executed within seconds.) In such a case the reputation value of the resource will be bounded with the θ-parameter of the h(u, r) function.

3 Malicious collectives with camouflage

This is a threat which is not always easy to tackle, since its resilience will mostly depend on the behavioural pattern followed by malicious peers (Gomez Marmol and Martınez Perez, 2009). These correspond to the malicious collectives with the variable behaviour. In our user reputation model, such variability could be partially detected with the SMUB model. Moreover, reputation value for such users will vary considerably over the time as well. Therefore, with such an approach it is possible to punish such behaviour with the reputation.

4 Malicious spies

This is a threat when malicious peers (spies) always provide good services when selected as service providers, but they also give the maximum rating values to those malicious peers who always provide bad services (Gomez Marmol and Martınez Perez, 2009). In Grids this corresponds to the situation when a user with high reputation provides the maximum rating to unreliable resources.

5 Sybil attack

An adversary initiates a large number of malicious peers in the system (Gomez Marmol and Martınez Perez, 2009). Each time one of the peers in the system is scheduled as a resource provider it provides malicious service and after that it is disconnected and replaced by another peer (Chakrabarti, 2007; Douceur, 2002; Gomez Marmol and Martınez Perez, 2009). In Grids, such an attack is hardly implemented in full form since appropriate certificate should be obtained from the certificate authority in order to integrate a resource into the Grid system. Other solutions to tackle this problem are to use the methodology of active experiment to monitor the availability of resources by sending, for example benchmark tasks, to incorporate Captcha mechanisms (von Ahn et al., 2008) or to require users to register with a valid telephone or credit card number (Kuhn et al., 2008). Up to this moment, there have not been many reports of Sybil attacks on Grid systems (Kuhn et al., 2008).

6 Man in the middle attack

A malicious peer can intercept the messages between other peers, rewrite the message and change reputation values (Gomez Marmol and Martınez Perez, 2009). Our model relies on the existing Grid security mechanisms to tackle this threat (Chakrabarti, 2007).

7 Driving down the reputation of a reliable peer

Malicious peers give the worst rating to those benevolent peers, who indeed provide good services (Gomez Marmol and Martınez Perez, 2009). Projecting onto the Grids, a malicious user will provide poor ratings to the resources, though user's jobs were completed successfully with appropriate QoS value. One possible way to tackle this problem is that jobs of users with low reputation value are never sent to resources with high reputation by the resource broker. Moreover, QoS metrics in Grids are measured by the monitoring system, and the malicious user should be able to illegally obtain necessary privileges to change these values and consequently the reputation value.

8 Partially malicious collectives

Malicious peers provide malicious actions for some kind of services, and, for others, they provide good services (Gomez Marmol and Martınez Perez, 2009). In Grids, this threat corresponds to users with variable behaviour (covered in malicious collectives with camouflage subsection), and to resources that for some types of services provide poor performance. By just considering a different score for every service offered by a resource, this threat is mitigated most of the times (Gomez Marmol and Martınez Perez, 2009). That is why, we included into our model a score function to provide different reputation values for different types of tasks executed by resources.

9 Malicious pre-trusted peers

Some models are based on the strategy that there is a set of peers that can be trusted before any transaction is carried out in the system, known as pre-trusted peers (Kamvar et al., 2003). This problem refers to assigning initial reputation to resources when a VO is formed. One possible way is to have a set of benchmarks tasks with the desired QoS metrics, execute them on all VO resources and assign the received reputation values to the resources.

Results of experiments

In this section, results of experiments are presented to assess the performance of the described model. The performance is evaluated in terms of resistance to security threat scenarios discussed in the previous section. A dedicated software application has been developed to run simulations with different scenarios.

1 Data description

In order to generate workload within experiments, i.e. jobs inter-arrival time and jobs execution time, we used real data provided by the Grid Observatory project[1]. This project provides data on job cycle in the EGEE grid infrastructure. In particular, we used data collected by the Real Time Monitor (RTM) systems that summarizes various information on jobs executed in the Grid. In total, the trace registers 37 attributes which categorized into Information, Timestamps and Metrics (Germain-Renaud et al., 2011).

2 Schedulers used in simulations

In all our simulations jobs are scheduled immediately after arrival. Two on-line schedulers were used in the study: a heuristic on-line scheduler that maps a job to a resource which provides the job earliest completion time (ECT), and a scheduler that uses resource reputation to map jobs (ECT-reputation). In order to integrate reputation into the latter scheduler a non-linear trade-off scheme (Voronin, 2011) is used.

Let ECT(ri) be the estimated completion time of running a job on resource ri, and ECTn(ri) is the corresponding normalized value:

$$ECT_n = \frac{ECT(r_i)}{ECT_{\max}}, \qquad (23)$$

where ECTmax is the upper bound value for the ECT value.

The ECT scheduler assigns job to a resource that minimizes the following expression

$$r^* = \arg\min_{r_i}\left(ECT(r_i)\right), \qquad (24)$$

while ECT-reputation scheduler assigns job to a resource that minimizes

$$r^* = \arg\min_{r_i}\left[\frac{\alpha_1}{1-ECT_n(r_i)} + \frac{\alpha_2}{rep(r_i)}\right]. \qquad (25)$$

where rep(ri) denotes reputation value of resource ri, and αk are parameters. In our experiments we used the following values for parameters: α1=α2=0.5.

3 Experimental parameters

---

[1] Grid Observatory: www.grid-observatory.org

All experiments were run for a Grid infrastructure of 20 resources with resource productivity (in unitless standard units) being uniformly selected from the range [1, 200]. Job complexity (also, in unitless standard units) was generated from traces provided by the Grid Observatory project lying in the range [1, 56000]. Distribution of job complexity is shown in Figure 1. Job execution time on a resource was estimated as jobComplexity/resourceProductivity. Jobs inter-arrival time and workload were also generated from EGEE traces. Figures 2 and 3 show cumulative number of submitted jobs over the time and job arrival rate (in jobs/min) respectively.

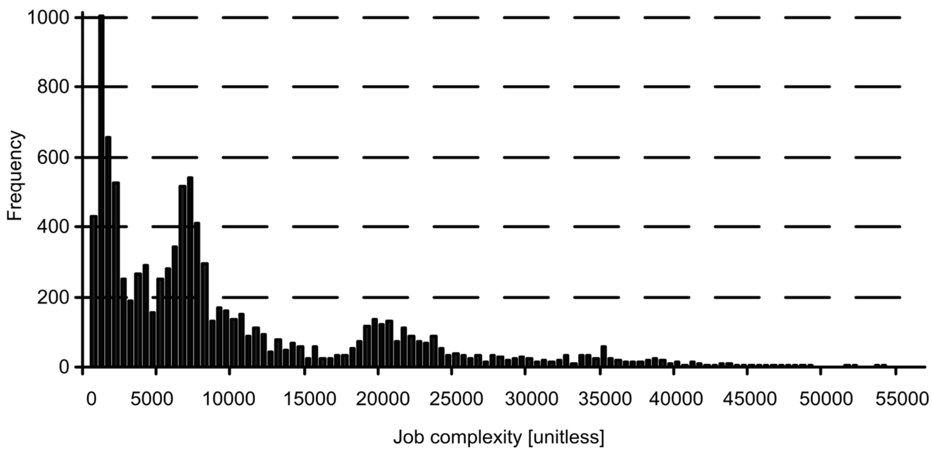


Figure 1. Distribution of job complexity within experiments (for 10000 jobs)

Within the experiments the following QoS metrics were considered: job waiting time, job execution time and job total completion time. The agreed SLA values were modelled as follows: the agreed waiting time was selected randomly from the range [1, 30000] sec, and the agreed execution time was selected as jobComplexity/minResourceProductivity. In order to simulate a scenario when a resource did not respect the agreed execution time the following approach was used: a random value from the interval [1, 2500] sec was added to the actual execution time value. The penalty function and reputation were estimated using Eq. (6) and (10) respectively. If not stated otherwise, the utility function (Eq. 5) was calculated for the job completion time QoS metric.

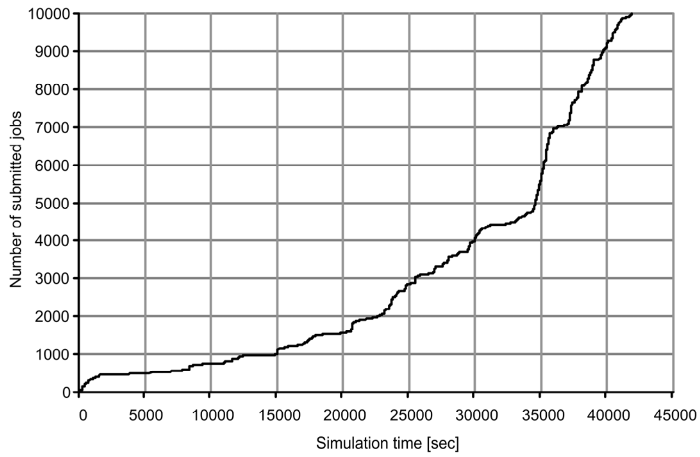


Figure 2. Cumulative number of submitted jobs

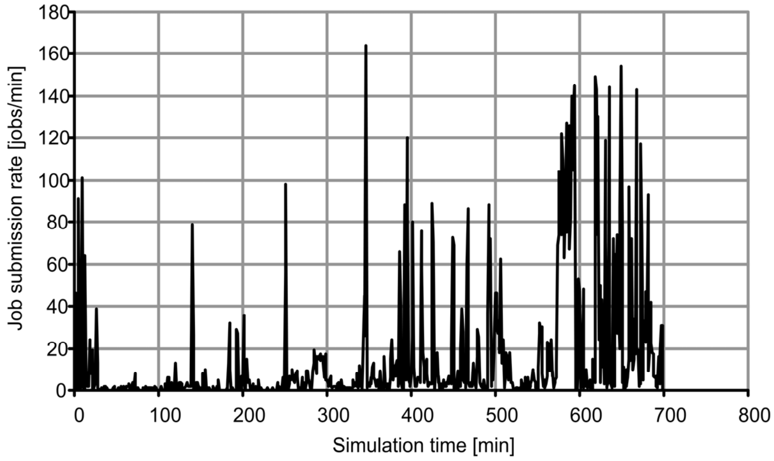


Figure 3. Job arrival rate

4 Analysis of security threat scenarios simulations

The following scenarios were run in order to evaluate the describe reputation model against security threat scenarios:

- Extreme cases. Some of the resources always provide bad services, and others always provide good services. Here, by “good” and “bad” services, we mean situations when a resource respects the agreed SLA, and when the agreed SLA is violated by a resource provider, respectively. This scenario also describes cases when a user tries to illegally improve reputation of the resource provider.

- Variable resources behaviour. Random and oscillating patterns of resource behaviour (Gomez Marmol and Martınez Perez, 2009) were considered within experiments. Within random pattern, at any time some of the resources provide either bad or good service. Within oscillating pattern, the resource is fully benevolent for a period of time and fully fraudulent for the next period, and so on (Gomez Marmol and Martınez Perez, 2009).

In both cases, ECT scheduler was applied as basic one.

Figure 4 shows how reputation for good and bad services changes over completion of the jobs.

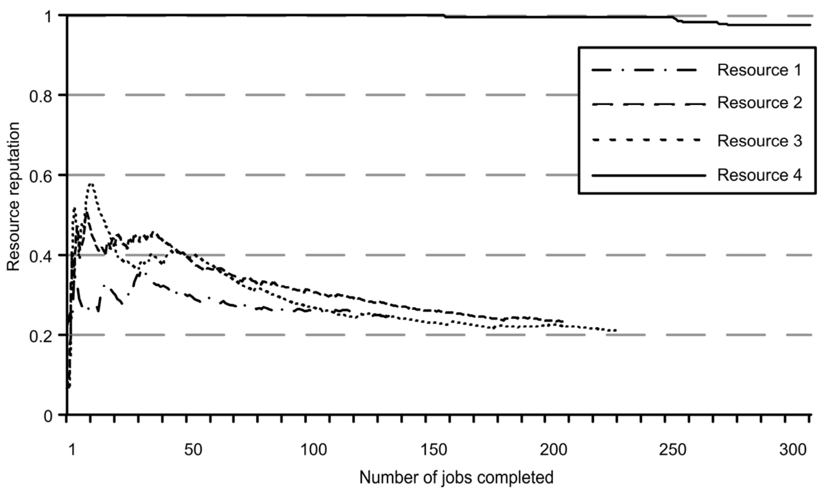


Figure 4. Reputation for resources that always provide good services (no. 1) and bad services (no. 2-4)

Since in our simulations random values are generated, it is important to analyse aggregated results for multiple runs. Figure 5 shows the average resulting reputation values at the end of simulations along with two standard deviations calculated for 10 runs.

The next scenario is when a user tries to illegally improve resource reputation. That behaviour will be captured by an alliance function (Eq. 15). Figure 6 shows an example when a user (from the same organisation as the resource) tries to illegally improve resource reputation by submitting a number of jobs for which the utility function is equal to alliance function h(u, r) (Eq. 15). Figure 6 shows how the resource reputation will change if no alliance function is applied, and if an alliance function is applied using different schemes for selecting θ-parameter (Eq. 15), in particular adaptive and fixed-value. From Figure 6, it is evident that adaptive scheme is preferable over the fixed-value scheme.

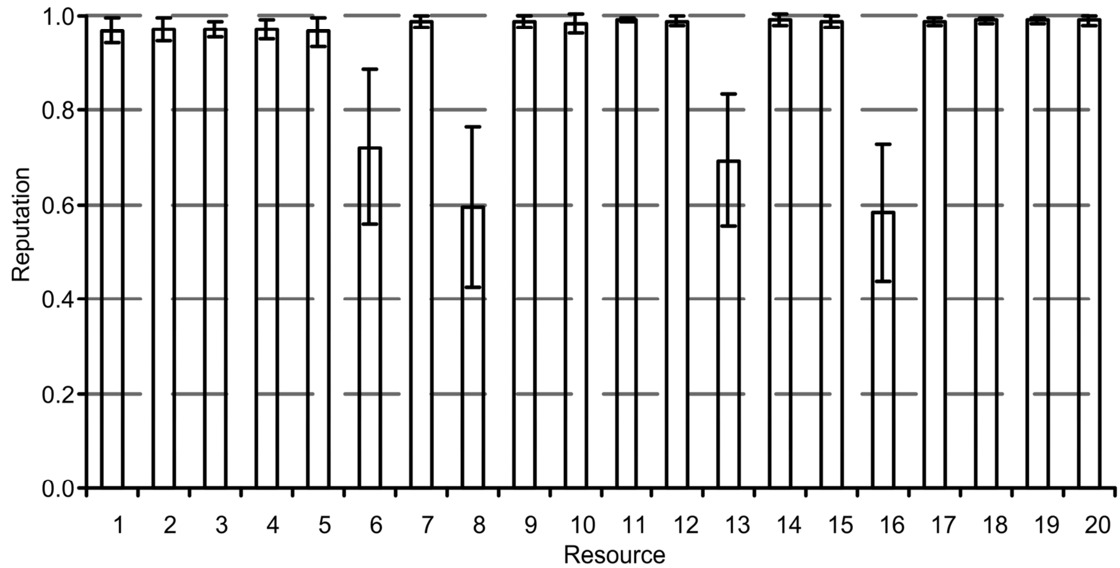


Figure 5. Resulting reputation averaged for multiple runs.
Resources with no. 6, 8, 13 and 16 always provide bad services

The following scenarios model the variable behaviour of resources. First, we consider random patterns that were modelled as follows: each "bad" resource was characterised by a trustworthiness rate. For example, if resource trustworthiness rate is equal to 0.6 then it meets the agreed SLA on average in 60% of cases. The following approach was used to simulate such scenarios: when untrustworthy resource was scheduled to execute a job, a random value uniformly distributed in the [0; 1] range was generated. If this random value was less than resource trustworthiness rate, then the resource met the agreed SLA. Otherwise, the agreed SLA is violated by the resource provider. Figure 7 shows the reputation of resources with random behavioural patterns with different trustworthiness rate.

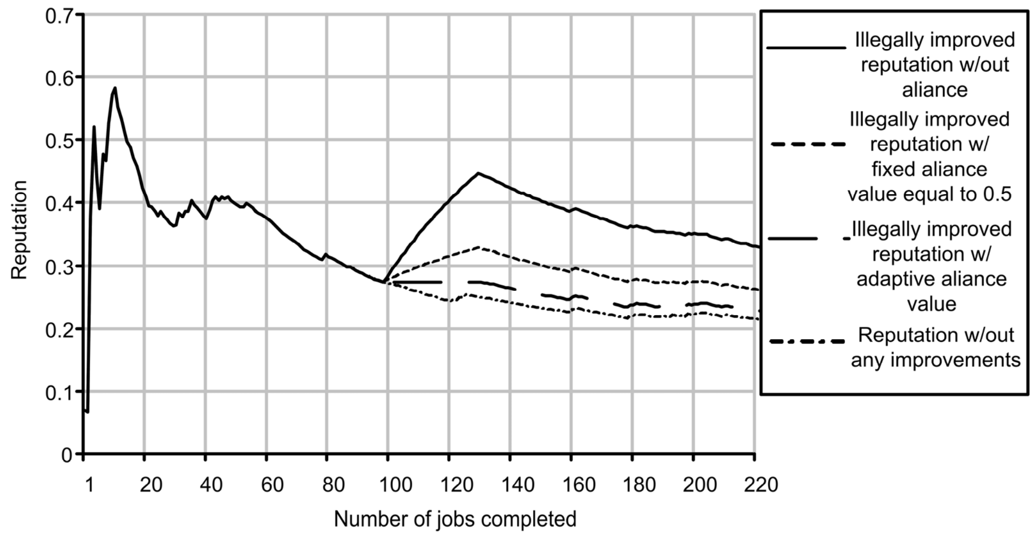


Figure 6. Situation when a user tries to illegally improve resource reputation. The user submits 30 jobs (no. 101-131) for which the estimated utility function (Eq. 17) is equal to the alliance function h(u, r) (Eq. 15). Without the alliance function, the utility would be equal to 1, and the reputation sharply increases.

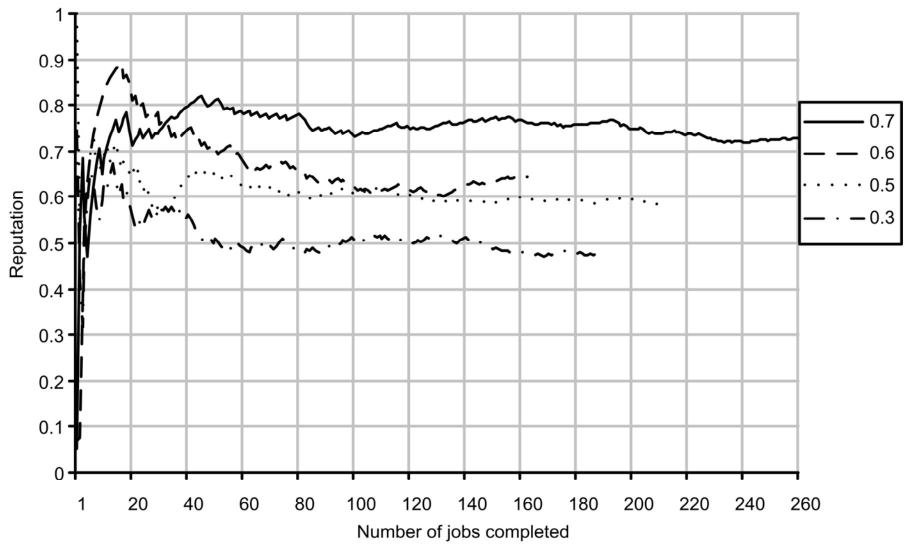


Figure 7. Scenario with resource random behavioural pattern at different trustworthiness rate

It is worth mentioning that the resulting resource reputation is close to trustworthiness rate (Table 2). It means that in such a case the proposed model was able to capture the variable pattern of resource behaviour.

Table 2. Comparison of trustworthiness rate and resulting resource reputation for random behavioural pattern scenario

Trustworthiness rate	Resulting resource reputation
0.7	0.728
0.6	0.644
0.5	0.582
0.3	0.475

The oscillating pattern suggests that the resource is fully benevolent for a period of time and fully fraudulent for the next period, and so on. Figure 8 shows how reputation changes for a resource with oscillating pattern.

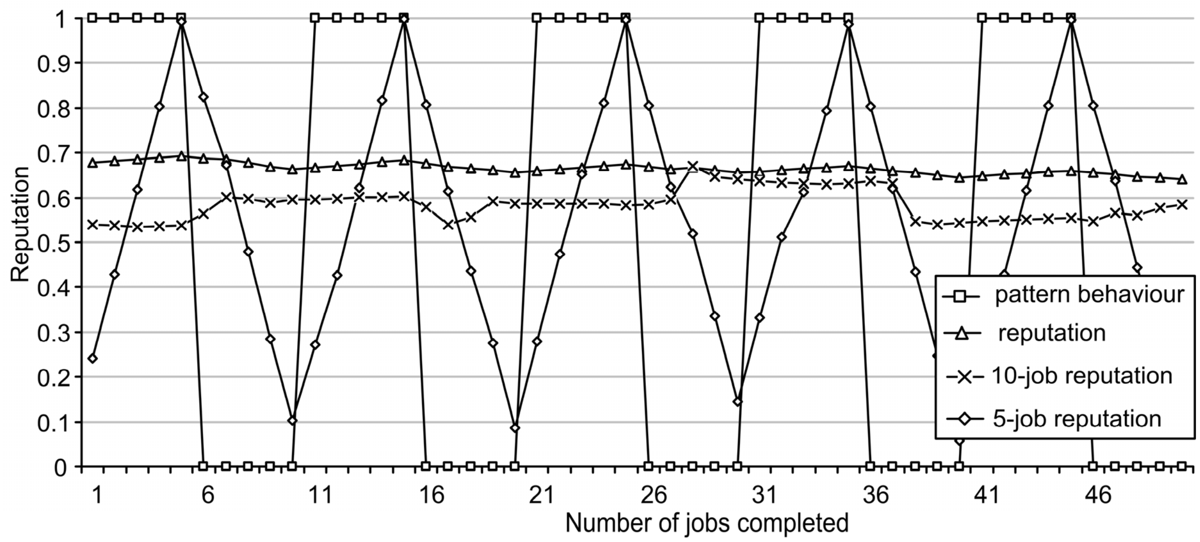


Figure 8. Scenario with resource oscillating behavioural pattern. In this example an oscillating period is equal to 5. Shown in the figure are: utility (Eq. 17), reputation estimated with no time decay function (Eq. 16), and reputation estimated with time decay function (for 5- and 10-jobs averaged utility)

From Figure 8, it is evident that just considering reputation without a time decay function does not allow us to detect the variable pattern of resource behaviour. It becomes partially possible when a parameter tc in Eq. (4) is appropriately chosen. In the shown example even a 10-jobs average utility does not exactly show the extent of variability.

The efficiency of the model was also tested against the error of the model (Gomez Marmol and Martınez Perez, 2009) which represents a malicious peers utilization (or selection percentage of malicious service providers). Here, experimental results are reported for both ECT and ECT-reputation schedulers. Within the first set of experiments we varied number of untrustworthy service providers (that always provide bad services). Figure 9 shows the error of the model depending on the number of untrustworthy resources.

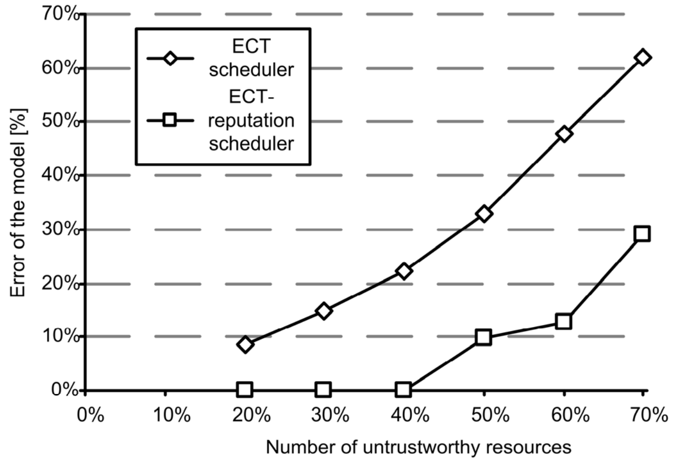


Figure 9. Error of the model depending on the number of untrustworthy resources

When using the ECT-reputation scheduler, the error of the model was below 15% when number of malicious resources was 60%, and below 30% in case of 70% of malicious resources. Within the second set of experiments we varied resource trustworthiness. We allowed 20% of the resources to be untrustworthy but with different degree of trustworthiness. Figure 10 shows the error of the model depending on resource trustworthiness rate.

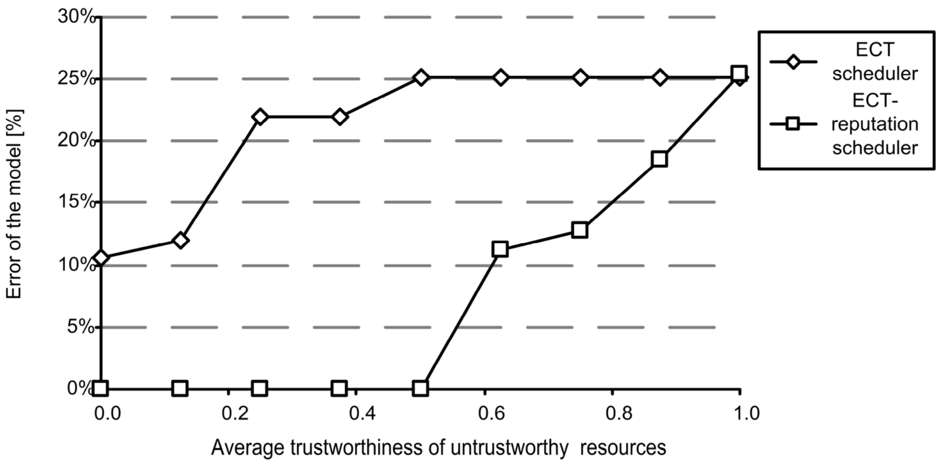


Figure 10. Error of the model depending on resource trustworthiness rate

The integration of reputation into the scheduler allowed us to reduce the error of the model. When using the ECT-reputation scheduler no jobs were scheduled until resource reputation became high (in our cases until average resource trustworthiness rate was more than 0.5, Figure 10).

Conclusions

In this paper we assessed most important and critical security threats for a utility-based reputation model in Grids. To tackle threats scenarios the model was further extended by incorporating a statistical model of user behaviour and several additional components, in particular: assigning initial reputation to a new entity in VO, capturing alliance between consumer and resource, determining time decay function, and score function.

The probabilistic estimation of trust model was proposed that allowed us to theoretically investigate properties of the model.

The experimental results showed that the model was effective in countering such threats as individual malicious peers, malicious collectives, driving down the reputation of a reliable peer. The error of the model was below 15% when number of malicious resources was 60%.

At the same time there were some limitations in countering malicious collectives with camouflage, in particular for oscillating behaviour pattern. Parameters in a time decay function have to be appropriately selected in order to detect the variable pattern of resource behaviour.

Future work should be directed on further investigation of oscillating patterns of resource behaviour and improving the model to counter these various patterns, and exploring application of the model for other large-scale service-oriented systems such as the Global Earth Observation System of Systems (GEOSS).

Acknowledgement

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Bibliography

[Abdul-Rahman and Hailes, 2000] Abdul-Rahman A, Hailes S. Supporting trust in virtual communities. In: Proc of the IEEE 33rd Hawaii Int Conf on Syst Sci, HICSS'00; 2000. p. 6007.

[Arenas et al, 2008] Arenas A, Aziz B, Silaghi GC. Reputation management in grid-based virtual organisations. In: Fernandez Medina E, Malek M, Hernando J, editors. Proc Int Conf on Secur and Cryptogr, SECRYPT 2008; 2008. p. 538–45.

[Azzedin and Maheswaran, 2002] Azzedin F, Maheswaran M. Integrating Trust into Grid Resource Management Systems. In: Int Conf on Parallel Process, ICPP 2002. Washington, DC, USA: IEEE Computer Society; 2002. p. 47–54.

[Chakrabarti, 2007] Chakrabarti A. Grid Computing Security. Berlin Heidelberg: Springer-Verlag; 2007.

[Douceur, 2002] Douceur JR. The Sybil attack. Lect Notes in Comput Sci 2002;2429:251–60.

[Eymann et al, 2008] Eymann T, König S, Matros R. A Framework for Trust and Reputation in Grid Environments. J of Grid Comput 2008;6(3);225-37.

[Foster et al, 2001] Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. Int J of Supercomput Appl 2001;15(3):200-22.

[Germain-Renaud et al, 2011] Germain-Renaud C, Cady A, Gauron P, Jouvin M, Loomis C, Martyniak J, et al. The Grid Observatory. In: 11th IEEE/ACM Int Symp on Cluster, Cloud and Grid Comput, CCGrid; 2011. p. 114–123.

[Marmol and Perez, 2009] Gomez Marmol F, Martınez Perez G. Security threats scenarios in trust and reputation models for distributed systems. Comput & Secur 2009;28:545–56.

[Marmol and Perez, 2008] Gomez Marmol F, Martınez Perez G. Providing trust in wireless sensor networks using a bio-inspired technique. In: Proc. of the Networking and Electronic Commerce Research Conf, NAEC'08, Lake Garda, Italy; Sep 2008.

[Haykin, 1999] Haykin S. Neural Networks: A Comprehensive Foundation. Upper Saddle River, New Jersey: Prentice Hall; 1999.

[Hluchy et al, 2010] Hluchy L, Kussul N, Shelestov A, Skakun S, Kravchenko O, Gripich Y, et al. The Data Fusion Grid Infrastructure: Project Objectives and Achievements. Comput and Inf 2010;29(2):319–34.

[Josang et al, 2007] Josang A, Ismail R, Boyd C. A Survey of Trust and Reputation Systems for Online Service Provision. Decis Support Syst 2007;43(2):618-44.

[Kamvar et al, 2003] Kamvar S, Schlosser M, Garcia-Molina H. The EigenTrust algorithm for reputation management in P2P networks. In: Proc of the 12th Int Conf on World Wide Web. New York, NY, USA: ACM Press; 2003. p. 640–51.

[Kerschbaum et al, 2006] Kerschbaum F, Haller J, Karabulut Y, Robinson P. PathTrust: A Trust-Based Reputation Service for Virtual Organization Formation. Lect. Notes in Comput Sci 2006;3986:193–205.

[Kuhn et al, 2008] Kuhn M, Schmid S, Wattenhofer R. Distributed asymmetric verification in computational grids. In: IEEE Int Symp on Parallel and Distrib Proces, IPDPS 2008; 2008. p. 1–10.

[Kussul et al, 2011] Kussul N, Shelestov A, Skakun S. Grid Technologies for Satellite Data Processing and Management within International Disaster Monitoring Projects. In: Fiore S, Aloisio G, editors. Grid and Cloud Database Management. Berlin Heidelberg: Springer-Verlag; 2011. p. 279–306.

[Kussul et al, 2009] Kussul N, Shelestov A, Skakun S. Grid and sensor web technologies for environmental monitoring. Earth Sci Inf 2009;2(1-2):37–51.

[Kussul and Skakun, 2004] Kussul N, Skakun S. Neural Network Approach for User Activity Monitoring in Computer Networks. In: Proc of Int Joint Conf on Neural Networks, Budapest, Hungary; 2004. p. 1557–62.

[Kussul and Novikov, 2009] Kussul O,Novikov O. Utility-based reputation model for VO in Grids. Visnuk NTUU "KPI": Informatics, control and computation 2009;50:137-145.

[Lecca et al, 2011] Lecca G, Petitdidier M, Hluchy L, Ivanovic M, Kussul N, Ray N, et al. Grid computing technology for hydrological applications. J of Hydrol 2011;403(1-2):186–99.

[Liang and Shi, 2010] Liang Z, Shi W. A reputation-driven scheduler for autonomic and sustainable resource sharing in Grid computing. J of Parallel and Distrib Comput 2010;70:111–25.

[Oh and Lee, 2003] Oh SH, Lee WS. An anomaly intrusion detection method by clustering normal user behavior. Comput & Secur 2003;22(7):596–612.

[Papaioannou and Stamoulis, 2008] Papaioannou TG, Stamoulis GD. Reputation-Based Estimation of Individual Performance in Grids. In: Eighth IEEE Int Symp on Cluster Comput and the Grid, CCGRID. Washington, DC, USA: IEEE Computer Society; 2008. p. 500–09.

[Schulter et al, 2008] Schulter A, Vieira K, Westphall C, Abderrahim S. Intrusion Detection for Computational Grids. In: Proc 2nd Int Conf New Technol, Mobil, and Secur. IEEE Press; 2008. p. 1–5.

[Shelestov et al, 2006] Shelestov A, Kussul N, Skakun S. Grid Technologies in Monitoring Systems Based on Satellite Data. J of Autom and Inf Sci 2006;38(3):69–80.

[Shelestov et al, 2008] Shelestov A, Skakun S, Kussul O. Intelligent Model of User Behavior in Distributed Systems. Int J on Inf Theory and Appl 2008;15(1):70–6.

[Shelestov et al, 2007] Shelestov A, Skakun S, Kussul O. Complex Neural Network Model of User Behavior in Distributed Systems. In: Proc of XIII-th Int Conf "Knowledge-Dialogue-Solutions", Varna, Bulgaria. Sofia, Bulgaria: FOI ITHEA; 2007. p. 42–9.

[Silaghi et al, 2007] Silaghi G, Arenas A, Silva L. A Utility-Based Reputation Model for Service-Oriented Computing. In: Priol T, Vanneschi M, editors. Toward Next Generation Grids. New York: Springer; 2007. p. 63–72.

[Skakun et al, 2005] Skakun S, Kussul N, Lobunets A. Implementation of the Neural Network Model of Users of Computer Systems on the Basis of Agent Technology. J of Autom and Inf Sci 2005;37(4):11–8.

[Song et al, 2005] Song S, Hwang K, Kwok YK. Trusted Grid Computing with Security Binding and Trust Integration. J of Grid Comput 2005;3(1-2):53-73.

[Srivatsa and Liu, 2006] Srivatsa M, Liu L. Securing decentralized reputation management using TrustGuard. J of Parallel and Distrib Comput 2006;66(9):1217–32.

[Sun et al, 2004] Sun HW, Lam KY, Chung SL, Gu M, Sun JG. Anomaly Detection in Grid Computing Based on Vector Quantization. Lect Notes in Comput Sci 2004;3251:883–6.

[Vieira et al, 2007] Vieira K, Schulter A, Westphall C, Westphall C. Intrusion Detection for Grid and Cloud Computing. IT Prof 2007;12(4):38–43.

[von Ahn et al, 2008] von Ahn L, Maurer B, McMillen C, Abraham D, Blum M. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. Sci 2008;321(5895):1465–8.

[von Laszewski et al, 2005] von Laszewski G, Alunkal B, Veljkovic I. Towards Reputable Grids. Scalable Comput: Pract and Exp 2005;6(3):95–106.

[Voronin, 2011] Voronin AN. A multicriteria problem of distribution of bounded resources. Cybern and Syst Anal 2011;47(3):490-3.

[Wu and Sun, 2010] Wu CC, Sun RY. An integrated security-aware job scheduling strategy for large-scale computational grids. Future Generation Comput Syst 2010;26:198-206.

Authors' Information

Olga Kussul – *Phd Student, National Technical University of Ukraine "Kyiv Polytechnic Institute", Prospekt Peremogy 37, Kyiv 03056, Ukraine; e-mail: olgakussul@gmail.com*

Major Fields of Scientific Research: Grid computing, information security, design of distributed software systems.

Nataliia Kussul – *Deputy Director, Space Research Institute NASU-NSAU, Glushkov Prospekt 40, build. 4/1, Kyiv 03680, Ukraine; e-mail: inform@ikd.kiev.ua*

Major Fields of Scientific Research: Grid technologies, design of distributed software systems, parallel computations, intelligent data processing methods, neural networks, satellite data processing, risk management and space weather.

Sergii Skakun – *Senior Scientist, Space Research Institute NASU-NSAU, Glushkov Prospekt 40, build. 4/1, Kyiv 03680, Ukraine; e-mail: serhiy.skakun@ikd.kiev.ua*

Major Fields of Scientific Research: Grid computing, Sensor Web, Earth observation, satellite data processing, risk analysis.

VIRTUAL EXPOSITIONS MODELING IN THE BULGARIAN FOLKLORE

Desislava Paneva-Marinova, Detelin Luchev, Konstantin Rangochev

Abstract: *In this paper we discuss the process of virtual exposition modeling for the folklore domain. Two folklore information systems are chosen as a source of the content – the Bulgarian folklore digital libraries and the Bulgarian folklore artery. They include specimens of different folklore narrative types (songs, rituals, faith, knowledge, proverbs, magic, etc.) and their audio-visual documentation. The purpose of the virtual expositions modeling is combining attractive, but difficult-to-collect objects that present completely the socium occupation, traditions and life in concrete (but not only) village (folklore region, area, etc.). It is important to present objects with different origin, for example juxtaposing available fund materials for the beginning of 20th century for one rite and this rite 100 year after. This approach will join the past with the present of the corresponding folklore objects and will present the Bulgarian folklore culture as vital and dynamically developing one.*

Keywords: *Virtual Exposition Modeling, Digital Libraries Services*

ACM Classification Keywords: *H.1 Models and Principles, H.3.7 Digital Libraries – Collection, Dissemination, System issues*

Introduction

Virtual expositions (VE) (also called exhibitions) for cultural heritage entail the bringing together of unlikely assemblages of people, things, ideas, texts, spaces, and different media. Curators, designers, artists, anthropologists, sponsors, visitors, artworks, artifacts, antiquities, machines, installations, display cases, spotlights, photographs, moving images, catalogues, promotional materials, object labels, audio tours, gallery guides – we might say that these constitute the apparatus of the exhibition experiment [Basu et al., 2007]

Virtual expositions modeling requires a big team work of specialists, clear idea, rich content repositories offering variety of objects and knowledge, design vision and technological infrastructure for exposition publishing. When we plan to create a temporary or permanent (thematic) exhibition of folk objects the team faces a series of non-trivial problem because of the specificity of ethnological knowledge and its hardly defined and fuzzy boundaries including types of folklore objects (tangible and intangible) and its complex structure.

In order to simplify the process of virtual exhibitions creation for the folklore domain several factors have to be considered: targeted applications and user groups, content repositories and its components, attractiveness and innovation of the exposition idea, relation with non-trivial facts, events, etc.

In this paper we discuss the process of VE modeling for the folklore domain. Two folklore information systems are chosen as a source of the content – the Bulgarian folklore digital libraries (http://folknow.math.bas.bg) and the Bulgarian folklore artery (http://folkartery.math.bas.bg) [Pavlov et al., 2011][Paneva-Marinova et al., 2010]. They are developed during the project "Knowledge Technologies for Creation of Digital Presentation and Significant

Repositories of Folklore Heritage"[1] project aiming to keep digital specimens from the Funds of the Institute of Ethnology and Folklore Studies with Ethnographic Museum at the Bulgarian Academy of Sciences.

The Bulgarian folklore digital library is a gallery of artefacts and knowledge for Bulgarian culture, art and folklore that will present a relatively limited number of specimens of different folklore narrative types (songs, rituals, faith, knowledge, proverbs, magic, etc.) and their audio-visual documentation. Until now, the Bulgarian folklore is always shown partially only with text, sound or image, but the authors' demand is for joint unities of words, music and motions. This possibility can be provided by contemporary multimedia environments. The ambitions of the authors are the demonstration of unique music dialects from different local folklore areas and advanced approaches for folklore content prescription representation through authentic sounds, videos, and photos of live rituals. Parts of the Bulgarian folklore specimens are presented from asynchronous point of view; other will be in their diachrony – unique materials, saved for years. Another task is the different record technique demonstration – inquiry, interview, inclusive observation, etc. It gives many ideas for virtual expositions and expects a wide range of potential users – professionals and scientists, non-professionals, connoisseurs and viewers, etc. [Paneva et al., 2007].

The Virtual Exposition Idea

When we speak about virtual exposition it is important to distinguish this concept from the digital collection. The difference between an exposition and a collection is that an exposition has a tight connection between its idea, objects, and script that ties them all together. It is this tight connection that is vital; otherwise, a virtual exhibition will "amount to little more than disorganized and decontextualized digital collections" [Silver, 1997].

As a key concept for one exposition is its idea. In defining the idea, we have to think about the different effects that we wish the exposition to create. Five types of effects that we may wish to consider are:

- Aesthetic: organized around the beauty of the objects
- Emotive: designed to illicit an emotion in the viewer
- Evocative: designed to create an atmosphere
- Didactic: constructed to teach about something specific
- Entertaining: presented just for fun

An aesthetic exposition is one that exists purely for the sake of presenting beautiful objects. Emotive exhibits exist primarily to elicit an emotion in the viewer. The purpose of an evocative exposition is to create a specific atmosphere for the viewer. Though it is hoped that the viewer will learn something from an exposition (be it virtual or gallery based), in some cases the exposition will have a specific didactic focus. Whether an exposition's purpose is to teach or to be a purely aesthetic experience, in nearly every case there is a level of pure

[1] The "Knowledge Technologies for Creation of Digital Presentation and Significant Repositories of Folklore Heritage" is a national research project of the Institute of Mathematics and Informatics, supported by National Science Fund of the Bulgarian Ministry of Education and Science under grant No IO-03/2006. Its main goal is to build a multimedia digital library with a set of various objects/collections (homogeneous and heterogeneous), selected from the fund of the Institute for Folklore of the Bulgarian Academy of Science. This research aims to correspond to the European and world requirements for such activities, and to be consistent with the specifics of the presented artefacts [Bogdanova et al., 2006].

entertainment. In many ways, it is the element of entertainment that separates an exhibition from a textbook or a lecture on Bulgarian folklore and ethnography. No doubt the objects of the Bulgarian folklore culture kept in the Bulgarian folklore digital libraries provides the abilities for reproducing these effects. For example, the collection “Goldsmiths and Jewels” represents the work of Bulgarian masters and from 18th to the early 20th century. The variety of jewels is correlated to the development of traditional Bulgarian costumes in different parts of the ethnic area. Bulgarian traditional jewels are made of similar materials and similar techniques of workmanship. Regional specific features and differences could be observed in shapes, ornaments and patterns. Casting, forging, hammering out, filigree, granulation are the traditional techniques applied by Bulgarian goldsmiths. Masters apply also enamel, engraving, hemstitch, mounting of stained glass, semiprecious stones and engraved mother-of-pearl plates.

The idea, or concept, behind an exhibition is what will set it apart from a random collection of objects. A few general topics for which every library or archive can find materials around which to build an online exhibition include the following:

- Anniversaries of births, deaths, or significant events in people’s lives
- Notable events in the life of an institution or region
- Specific materials from certain collections or subcollections
- Themes built around materials in the collection
- Treasures
- Work done by various departments of the library, archives, or other units or departments of the parent institution
- Odd and unusual [Kalfatovic, 2002]

Each of these areas can be tailored and focused to reflect the strengths of an individual library or archive. We may look at different examples of the exhibitions in the National Ethnographic museum in Sofia (http://www.eim-bas.com/museum.php?p=exhibitions&l=en): “Jubilee Exhibition of the Artist Evgenia Lepavtsova” had dedicated 50 years of her life to ethnography with her unique ethnographic picture, drawings and illustrations (anniversary), “The Holy Path – the Life of Bulgarian Jews”, jointly exhibition with the Museum of Sofia, Central State Archives, the Museum at Sofia Synagogue was dedicated to the 50th anniversary of Israel state (notable events), “Bulgarian Folk Costumes” (specific materials from certain collections or subcollections), “Treasured for the Generations”, etc.

Moreover, when we prepare a Bulgarian folklore exposition it is important to have in mind that objects included could be as simple as complex. Example of a complex folklore object is CFO A1_146_2-14, an interview containing information of the catholic community in the village of Oresh, Svishtov region, northern Bulgaria (see figure 1) [Paneva-Marinova et al., 2010]. The emphasis in the interview is on the ritual, festival, and everyday life in the village, on the popular beliefs and knowledge. Every one of these folklore object types also has several sub-categories, depicted on figure 1.

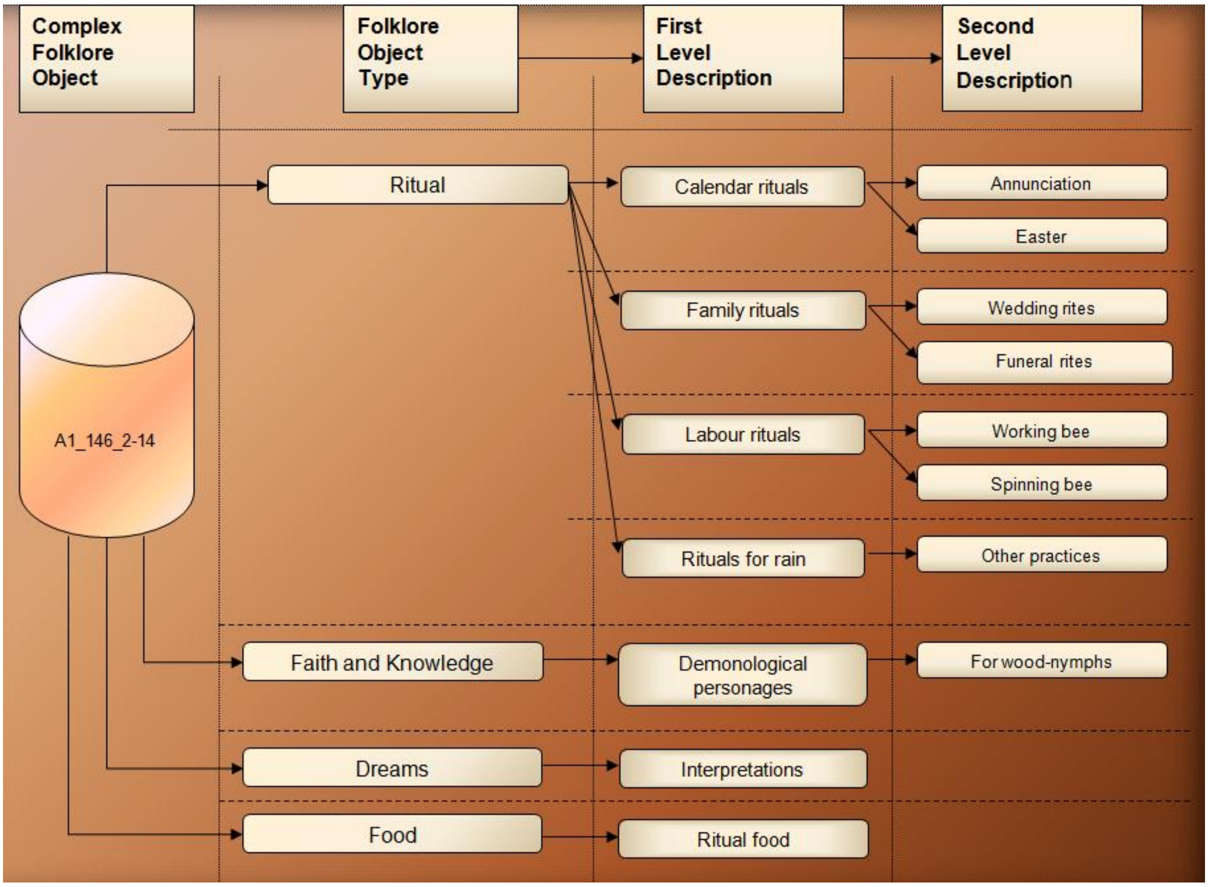


Figure 1: Example of a complex folklore object

The Virtual Exposition Planning Process

The exposition planning process is composed of a number of distinct steps:

- Preparation of the exposition proposal
- Proposal evaluation
- Selection of objects
- Drafting of the script
- Preparation of objects
- Exhibition design and Web creation
- Final editing
- Additions, changes, corrections

As with gallery exhibitions, there are numerous possibilities for organizing your idea and objects. Among these possibilities are:

- Object-oriented organization
- Systematic organization
- Thematic organization
- Organization by material type
- Organization by multiple schemes

As a typical object-oriented exposition in Bulgarian ethnology and folklore could be specified "the magic mask". The systematic organization of "Bulgarian folk costumes" exhibition follows the idea of presentation of different materials, techniques, regional specifics, etc. An interesting idea for thematic organization is the show of Christian images and symbols in traditional Bulgarian jewellery. Organization by material type of the objects - cooper, wood, clay, etc., gives possibility not only for separate, but for parallel expositions on the base of motifs, techniques, etc.

Important steps are selecting and preparing of the objects. Most likely reasons for selecting an object for exhibition are that, in the opinion of the curator, the object is intrinsically of interest, or information about it is considered of value to the visitor, or the object has a contribution to make to a more general story which the visitor is to be told [Belcher, 1991]

Virtual Expositions Management

The virtual expositions management will cover the basic processes on creating, preview, update and close exposition [Paneva-Marinova et al. 2011] (see Figure 2). The process will be executed using the digitized content and artifacts from the Bulgarian folklore portal (connected to the Bulgarian folklore digital library).

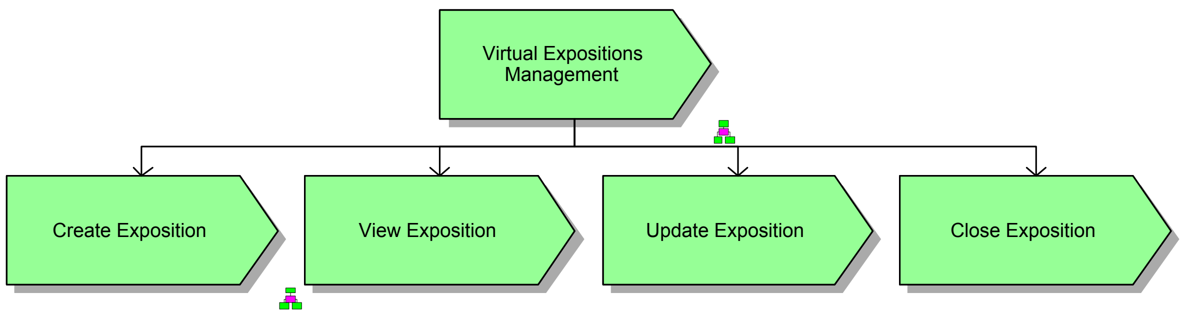


Figure 2: VACD of the virtual expositions management process

Figure 3 depicts an EPC diagram of the "Virtual expositions management" process. It shows the control flow structure of the process as a chain of events and functions. The functions present the actions and the tasks that must be implemented as a part of a business process, e.g., discussion of a proposal, build an exposition query, etc. Usually the functions add extra value to the process. The functions have input resources (e.g., documents), create output results (e.g., the “launch exposition” function creates an exposition) and could spend a resource (e.g., human). The events constitute the changing state of the world after the execution of a process, e.g., a query created, exposition objects selected, etc. The events described the situation before and after an action are executed. The functions are linked to events by logical connections. In this way the control flow is defined [Davis, 2001].

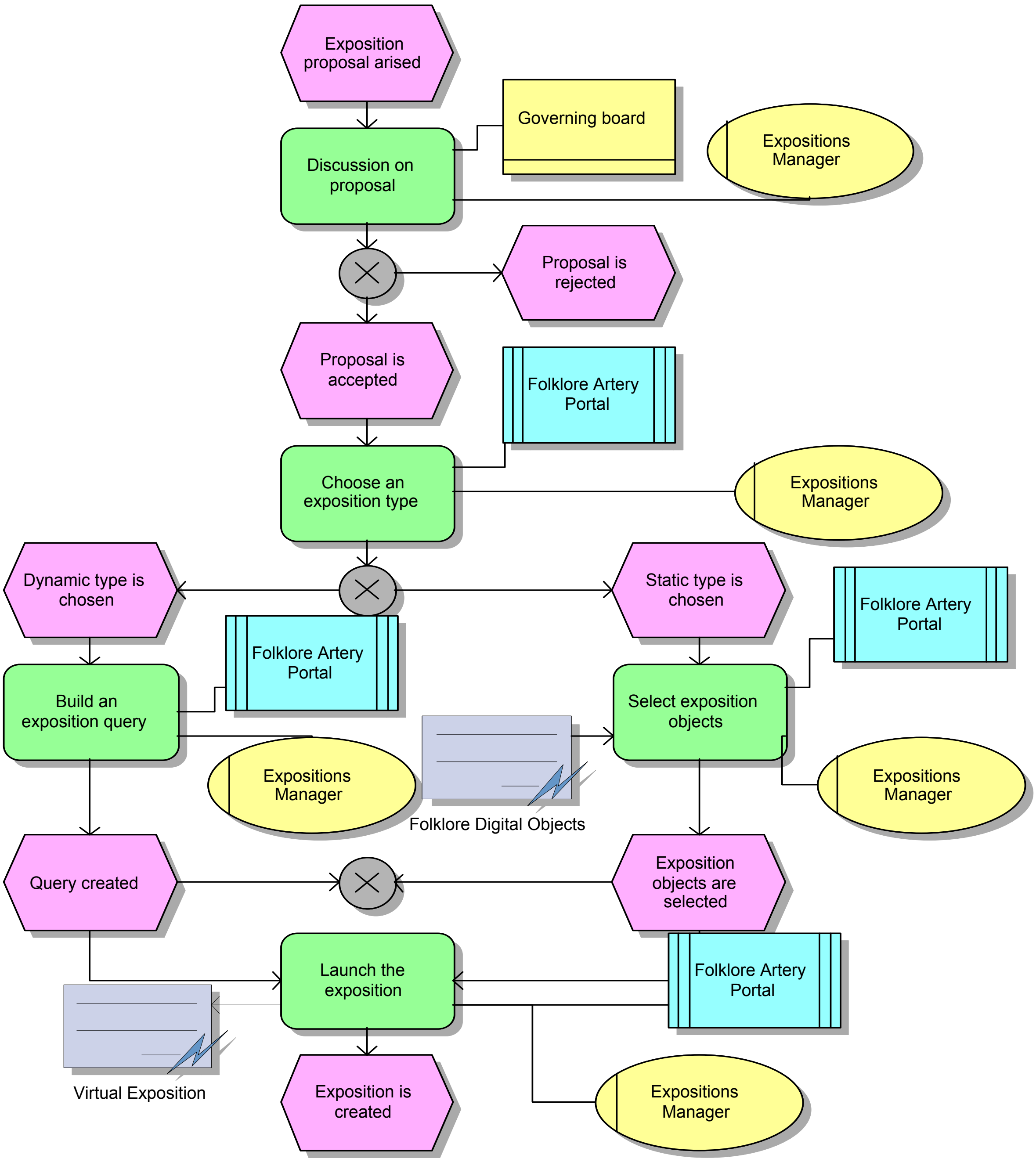


Figure 3: EPC diagram of the virtual expositions management process

Conclusion and Future Work

Nowadays, online exhibitions are a regular offering from the cultural institutions. They have also become an almost necessary adjunct to traditional physical exhibitions, offering a continuing life to the ideas presented. Additionally, we are seeing an increasing number of virtual-only exhibitions in which memory institutions are using the traditional notions of the exhibition as springboards to create interesting, instructive, and fun exhibitions that will never see visitors walking through them. The benefits of online exhibitions are many and include the abilities to showcase objects that could never be on view in a gallery space due to their fragility and value or to present life human traditions, far less expensive than a festival organization, to exhibit collection of objects, changeable in time (for example VE of Bulgarian folklore calendar rituals, presented through photos, films, sketches, technologies description, etc., the exhibitions of Easter eggs occurs two weeks before Easter; Christmas rituals, songs, labels – in December, etc.). As we move to provide increased access to folklore collections in the form of online exhibitions, it is important to remember why we are creating expositions. The purpose of the virtual expositions modeling is combining attractive, but difficult-to-collect objects that present completely the socium occupation, traditions and life in concrete (but not only) village (folklore region, area, etc.). It is important to present objects with different origin, for example juxtaposing available fund materials for the beginning of 20th century for one rite and this rite 100 year after. This approach will join the past with the present of the corresponding folklore objects and will present the Bulgarian folklore culture as vital and dynamically developing one.

Bibliography

[Basu et al., 2007] Basu, P., S. Macdonald. Introduction: Experiments in Exhibition, Ethnography, Art and Science", in Basu, P. and Macdonald, S (eds.), 2007, Exhibition Experiments. Oxford, Blackwell.

[Belcher, 1991] Belcher, M. Exhibitions in Museums. Washington, D.C., Smithsonian Institution Press, 1991.

[Bogdanova et al., 2006] Bogdanova, G., R. Pavlov, G. Todorov, V. Mateeva. Technologies for Creation of Digital Presentation and Significant Repositories of Folklore Heritage, Advances in Bulgarian Science Knowledge, National Center for Information and Documentation, 2006, Vol. 3, pp. 7-15.

[Davis, 2001] Davis, R. Business Process Modeling with ARIS: a Practical Guide, Springer, 2001.

[Kalfatovic, 2002] Kalfatovic, M. Creating a Winning Online Exhibition: A Guide for Libraries, Archives, and Museums, Chicago and London, 2002.

[Paneva et al., 2007] Paneva, D., K. Rangochev, D. Luchev. Ontological Model of the Knowledge in Folklore Digital Library, In the Proceedings of the Fifth HUBUSKA Open Workshop "Knowledge Technologies and Applications", 31 May - 1 June, 2007, Kosice, Slovakia, pp. 47-55.

[Paneva-Marinova et al., 2010] Paneva-Marinova, D., R. Pavlov, K. Rangochev. Digital Library for Bulgarian Traditional Culture and Folklore, In the Proceedings of the 3rd International Conference dedicated on Digital Heritage (EuroMed 2010), 8-13 November 2010, Lymassol, Cyprus, pp. 167-172, Published by ARCHAEOLINGUA.

[Paneva-Marinova et al. 2011] Paneva-Marinova, D., R. Pavlov, M. Goynov. Business Modeling of the Application Architecture of the Bulgarian Folklore Artery, In the Proceedings of the First International Conference "Digital

Preservation and Presentation of Cultural and Scientific Heritage, September 11-14, 2011, Veliko Tarnovo, Bulgaria, pp. 43-50.

[Pavlov et al., 2011] Pavlov, R., D. Paneva-Marinova. Digital Libraries and Portals Saving National Cultural Heritage (IMI-BAS Experience) (Invited Talk) In the Proceedings of the First International Conference “Digital Preservation and Presentation of Cultural and Scientific Heritage, September 11-14, 2011, Veliko Tarnovo, Bulgaria, pp. 182.

[Silver, 1997] Silver, D. Interfacing American Culture: The Perils and Potentials of Virtual Exhibitions, American Quarterly, 1997, 49, no. 4, pp. 825–850.

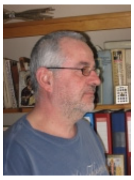
Authors' Information

***Desislava Paneva-Marinova** – PhD in Informatics, Assistant Professor, Institute of Mathematics and Informatics, BAS, 8, Acad. G. Bonchev Str., Sofia 1113, Bulgaria; e-mail: dessi@cc.bas.bg*

Major Fields of Scientific Research: Multimedia Digital Libraries, Personalization and Content Adaptation, eLearning Systems and Standards, Knowledge Technologies and Applications.

***Detelin Luchev** – PhD in Ethnology, MA in Informatics, MA in History, e-mail: luchev_detelin@abv.bg*

Major Fields of Scientific Research: Communities and Identities, Ethno-statistics, Museums and Archives, Digital Libraries.



***Konstantin Rangochev** – PhD in Philology, Assistant Professor, Institute of Mathematics and Informatics, BAS, 8, Acad. G. Bonchev Str., Sofia 1113, Bulgaria; e-mail: krangochev@yahoo.com*

Major Fields of Scientific Research: Ethnology, Folklore studies, Culture Anthropology, Linguistics, Computational Linguistics, Digital Libraries.

MODEL FOR IT TRAINING AND EMPLOYMENT OF PEOPLE WITH AUTISM SPECTRUM DISORDERS

Ekaterina Detcheva, Mirena Velkova, Ani Andonova

***Abstract**: ESI (European Software Institute), Center Eastern Europe and BASSCOM, in collaboration with Association Autism developed a project of a model for employment provision to people with ASD. The model includes trainings and workshops for IT companies for work with people with ASD, as well as theoretical/ practical IT training for the job candidates.*

The job positions for the employees with autism spectrum disorders are software products testing for bugs, data processing in IT systems, administration and office functions and other suitable activities using IT.

This paper describes the course of the project and the sustainable results of the pilot model. The training program and the methodology for adapted and real employment, developed within the framework of the project are also described.

***Keywords**: IT training and employment, autism, social iInclusion.*

***ACM Classification Keywords**: J. Computer Applications - J.4 Social and Behavioral Sciences, K.3 Computers and Education - Assistive technologies for persons with disabilities, K.3 Computers and Education - Employment, K.3 Computers and Education - Handicapped persons/special needs,*

Introduction

The project “Model for IT training and employment of people with ASD (Autism Spectrum Disorders)” is developed in 2007 and expresses the willingness of IT companies – members of BASSCOM (Bulgarian Association of Software Companies) to provide employment for people with disabilities. In April 2011 the project implementation started focusing on the training and employment of people with ASD.

ESI (European Software Institute) Center Eastern Europe and BASSCOM, in collaboration with Association Autism developed a project of a model for employment provision to people with ASD. The model includes trainings and workshops for IT companies for work with people with ASD, as well as theoretical/ practical IT training for the job candidates.

The job positions for the employees with autism spectrum disorders are software products testing for bugs, data processing in IT systems, administration and office functions and other suitable activities using IT.

Why information technologies? The work with IT offers a safe and foreseeable (logically managed) working and training environment for people with autism spectrum disorders. IT are a way for overcoming of the social isolation and helps the adequate social integration of people with ASD.

In July 2010, the program PROGRESS of EC in the area of Equal Rights and Social Inclusion announce a call for project proposals for co-financing of pilot projects on employment for people with autism spectrum disorders. More than 50 projects from all the countries, members of the EU, are submitted. Only four of them are approved for financial support and among them is that of ESI Center Eastern Europe, BASSCOM and Association Autism.

On the 18th of April 2011 the implementation of the project "Development and Piloting a Model for Occupational Training and Employment of People with ASD in the ICT Sector" started its implementation, co-financed by the European Union on the program "PILOT PROJECTS ON EMPLOYMENT OF PERSONS WITH AUTISM SPECTRUM DISORDERS".

The project duration is 12 months. It is focused on improvement the conditions for employability of the people with ASD through development and piloting of sustainable employment model for adjusted vocational training and individual support, which are to compensate the deficit, caused by the impairment.

Partners

ESI (European Software Institute) Center Eastern Europe – Project Coordinator. ESI CEE supports IT companies and organizations in the implementation of leading strategic management and software engineering methodologies. The mission of ESI CEE is to increase industry capacity and productivity in the region, ICT business competitiveness, and ICT professional skills and qualifications.

BASSCOM (Bulgarian Association of Software Companies) – organizes the contacts and activities related to ICT companies. The main goals of BASSCOM are to promote the Bulgarian software industry, develop the professionalism and competitiveness of IT sector, work for improvement of education system and participate actively in development and implementation of effective IT policies in benefit of the entire society.

Association Autism – organizes the activities related with training and employment of persons with ASD. Association Autism is a parental and social organization which supports people with ASD. The mission of the Association is to support the effective social integration of people with ASD.

Estimated Results

The goal of this project was to improve the employment situation of people with ASD through development and piloting of a sustainable employment model, tailored vocational training and individualised support to compensate the deficit caused by the impairment.

The project partners developed and pilot an employment model for provision of training and employment to people with ASD. A tailored ICT training and certification on IT skills for trainees with ASD, a wide awareness campaign towards employers and further employment to people with ASD in ICT companies were implemented within the project.

The project goal was be achieved through a set of activities related with research of the present situation of people with ASD in the labor market, possible preventive measures ought to be taken in order to prevent unemployment and exclusion of people with autism and Asperger Syndrome from the labor market, training needs analysis, adaptation of an existing ICT training curriculum for people with ASD to the needs of people with ASD, where appropriate, conducting ICT trainings and certification of ICT skills, training tutors to work with people with ASD, provision of employment in ICT companies, raising the awareness in the community and verification/ dissemination of the project results. An approach based on individual assessment and development of the potential and capabilities of people with ASD was used. The trainings itselves were focused on the quality and productivity in the work process and hereby demonstrated the assets for companies which hire people with disabilities. Raising the awareness on the advantages from hiring people with ASD and their abilities was another direct approach for the involvement of more employers from the ICT sector and thus ensuring a sustainable employment rates for people with ASD.

The project consortium included an ICT professional training organization, service provider association of people with ASD and an ICT branch organization. The initiative expressed the willingness of IT companies – some of them members of BASSCOM (Bulgarian Association of Software Companies) to provide employment for persons with autism in the IT sector. The project is co-funded by EC. Young persons with autism and potential employers were trained how to work together in real business environment. After the accomplishment of the trainings the youths have been provided with employment in Bulgarian ICT intensive companies. The trainings consisted of workshops for the employers on how to work with persons with autism, IT trainings for persons with ASD – theoretical and practical IT trainings. The training of the companies targeted employers and computer skills trainers to help them understand the social model of the disability and learn how to handle employees or trainees within the autistic spectrum. It presented techniques that challenged the traditional perceptions of disability and ASD, transfered knowledge for communication with disabled people, gave clues for supportive company culture, provided examples of successful good practicies from other countries. This training was delivered by a team of experienced experts in the field of autism. All persons with autism and experts dealing with the integration and rehabilitation of people with ASD passed through IT card test for personal certification of computer literacy. The project idea was to provide the people with autism the opportunity to move away from caring home environment during the day, find a way for professional and personal development with means of information technologies, and work in real business environment. The companies which took part in the project and hired persons with autism would expect improvement of the working atmosphere, positive changes in the employees' perception towards surrounding environment – through pro-activity in the community interaction to pro-activity in organization goals achievement.

Methodology

During the first half of the project implementation 9 IT companies have been starting to prepare for hiring the youths. This phase of the project included The Disability Equality Training for Employers as a one day course for managers and human resources professionals from potential beneficiary organisations/companies. The purpose of the training was to help employers gain a better understanding of the disability legislation and how it impacts hiring, supervising and working with people with ASD. The training was also focused on raising the participants' awareness of the employment issues of disabled people – psychological characteristics, behavioural specifics, work tasks considering the skills of people with ASD etc. Recognising how policies and practice can be easily changed to create a welcoming environment at the workplace for the ASD employees was one of the outcomes from the training. It created an understanding what they can do to remove barriers to employment and promotion of ASD staff. The following training was for The Disability Equality in the Workplace. It involved peers, middle management and logistics management from the potential beneficiary organisations /companies. This training provided practical guidelines for inclusion of employees with ASD. The idea of that was to equip the staff member with a take-away pack of useful reference materials as well as encourage them to participate in the process of integration of the people with ASD within their team at the work place. This trainings were conducted individually in each company.

By the end of the project 7 companies actually created a work opportunity for the autistic people. They provided internship for job positions related with IT apps functional testing, content management systems, digitalization and processing of documents, data base filling, and company social activities.

The following step took care of the evaluation of the 20 candidates and the preparation of the individual plans of the participants. The characteristics of each person were taken into consideration with regards to the learning process, personal level of experience in working with computers, ability to study in a group, the speed of acquiring new skills and competences, the topics of motivation and the points of interest etc.

The computer training phase was planned to be up to 4 hours a day. Due to the lack of similar commitments of the participants it started with an hour a day and grew till 4. The whole training was generally divided into four modules and each module cover different set of disciplines. The first two included the MS & Open Office applications: (e.g. correspondence information technologies), Word, Windows XP, MS Office XP (Word, Access, Excel, Power Point); Spreadsheets (MS Excel, Open Office Calc), Internet – applications (e.g. e-mail, address books, browsing, work with search engines, e-commerce, e-signature, finding information and software in the Internet). The trainees' advancement in each discipline was monitored through regular tests and upon completion participants took an exam, the results from which defined the selection of trainees for the next training module/phase. There were no participants who advanced so significantly so they can cover the Object-oriented programming, data structures; Data bases and logical programming, applications – e.g. Java or any part of the aadvanced and functional programming, design and analysis of algorithms, artificial intelligence. After the process of computer training was completed the psychologists and special teachers from the professional team of experts hold another round of meetings with the companies representatives so to be able to chose the most appropriate applicants for the positions offered by the companies to the people within ASD.

Prior to the beginning of the internship the companies were equipped with a detailed handout how to handle and work with a person with ASD. With addition to that they received an individual plan and instruction guide how to particularly work with the person to be invited to their company.

The start of the internships for the young people was secured by the presence of an ASD expert to support the mentors and the teams of each IT company which offered a workplace for the autistic participants. The professional psychologists who were responsible for the training of the particular people with ASD were the experts who accompanied the interns at the workplace for the first few days of the new phase of their further training. For the period of the internship there was an ongoing phone and e-mail communication between the ASD specialists and the mentors in order to provide the best possible environment for the integration of persons with ASD.

The challenges met by the people with autism were during the process of training were with regards to the new area of expertise which they didn't have much of experience with other then playing and using it for purposes different than work. The requirements for the level of independency were pretty high and the influence support or interference of the family members in the process reduced to a minimum. There was an awareness of responsibility and obligations to the tasks in the workplace required. Some of the participants never had the chance to be previously engaged in a long term commitment other than school and play. The difficulties in communication were also among the challenges experience by the young people with ASD. There was a new environment, new rules, new people and atmosphere they needed to get used too. The accomplishment of the undertaking required a certain level of task prioritization and time management which the project participants had to build or improve.

Other difficulties to tackle for the ASD people were related to the need to have a better ability to transition from English to Bulgarian and then English again while proceeding with the filing and data taking. The companies' teams were really tolerant to the specifics of the ASD colleagues but the lack of interaction made them wonder and not be sure if they created a welcoming enough environment for their autistic interns. In some of the cases the attitude of the ASD people was too direct or too loose and the talking in a loud voice was disruptive.

Conclusion

Despite of the challenges there are obvious results to report. The main achievements are with regards to the new acquired computer skills, newly created interest in the IT matter and computer supported communication. The improvement of social interaction skills during the internship even before the completion of the project was significant. It impressed even the ASD experts and family members of the participants. The positive development of the social skills could be witnessed in the process of communication with work colleagues but also in the relationship with the extended family and the outside world.

The basic IT knowledge and competences learned will increase the employability of the participants in the future if they will not be hired by the end of this project. 6 IT companies associated with BASCOM/ESI Centre Bulgaria provided the necessary accommodations and environment adaptations for each ASD individual on the job and know more about handling disabled staff members following the disability equality trainings provided to their middle management and HR officers. The IT people who took the roles of mentors and those who were closely helping them also gained a significant amount of knowledge to communicate better with disabled trainees in their practice. While preparing themselves for the computer training of the ASD participants 10 special pedagogues and psychologists improved and structured their previous IT skills and knowledge which can be also seen as a positive effect of the project undertaking. The experience learned will be used in the follow-up stage when the model of education and employment support will be applied at a larger scale.

The review of the employment of ASD persons will give us a chance to recommend legislative changes that will improve the effectiveness of the public funds and the employment situation of more disabled people. The difficulties in this area are with regards to the higher percentage of disability the people with ASD usually get. At the same time by issuing the medical papers with 70% of disability does not give the chance to autistic people to be hired. The controversy of this situation is rather complicated. There is still a lack of understanding of the autistic spectrum disorders. It is difficult to accept that a person with 70% and more percent of disability who even didn't graduate from school could be very advanced professional when it comes to the field of computers.

The positive results of the integration of the ASD participants in the 6 companies who choose to cooperate with ESI Centre Eastern Europe and Association Autism in this project could be seen a positive sign for future application of this model throughout Bulgaria and the neighbouring countries. As more adults with autism are entering the workforce than ever before, the issues involving autistic people and the workplace are being redefined to benefit both employees and employers. The creation of proper handout materials explaining what autistic people expect in the workplace and the type of resources available to help achieve a successful employment experience is adding to the future opportunities in the process of recruitment of ASD people into the free market.

Just a few years ago the job market was slim or almost closed for people with autism. Opportunities were scant due to factors such as less awareness about autism, stereotypes about the disorder and communication problems during interviews with potential employers. Today autism awareness around the world and support services for people with autism have led to more opportunities for autistic adults in every level of the job market. More employers are aware that many people with autism have valuable skills and qualifications that can benefit their companies in different areas of the economy. The projects' final conference was held on 13th of March 2012. There were Belgium and Spanish representatives sharing experience of other EU countries with regards to successful good practices in recruiting autistic people and using IT service for the purpose of quality of life support and career opportunity. The projects presented at the Sofia conference were used to convince the decision makers in Bulgaria as well as the different businesses that our initiative is not an isolated phenomenon. The reason to let government officials and conference participants hear about the experience in EU countries in the field was to convince them of the need to create a better environment for busting the integration of people within ASD into the free labour market. The completion of this project proved that even people with ASD with no previous computer experience can do useful tasks within a company. Depends on the fact whether the person is low or high-functioning, an autistic adult can excel at tasks involving categorization, such as packing goods at a shop floor, or solving high- tech software problems. Knowing and understanding more and more the autistic spectrum disorder allows us to take action and turn the specific characteristics of autistic people into their advantage when it comes to work commitment.

This exemplar is meant to be a model for manuscript format. Please make your manuscript look as much like this exemplar as possible. In case of serious deviations from the format, the paper will be returned for reformatting.

NOTE: Every manuscript submitted to be published by ITHEA ISS need to be presented by author(s) personally at any of the ITHEA International Conferences.

The Project in the News

http://bnt.bg/bg/news/view/59330/s_malko_pomosht_ot_prijateli_za_horata_autisti_u_nas

ESI CEE – Employers workshop BNT 1

http://bnt.bg/bg/news/view/59288/fandykova_otkriva_zona_za_deca_autisti

Official opening of renovated facilities in Center for Social Rehabilitation and Integration of people with ASD BNT 1

http://www.btv.bg/story/428248679-Nadejda_za_horata_s_autizam.html

Official opening of renovated facilities in Center for Social Rehabilitation and Integration of people with ASD bTV

http://bnt.bg/bg/news/view/71835/softuerni_kompanii_naemat_mladeji_s_autizym

National Television Broadcast - March 2012

http://www.karieri.bg/karieri/novini/1786192_autistite_imat_shans_za_kariera_v_it_sferata/

Careers newspaper, 13 March 2012

http://esicenter.bg/news.aspx?nid=48

Acknowledgement

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Authors' Information

Ekaterina Detcheva *– Institute of Mathematics and Informatics, BAS, Block 8 Acad. G. Bonchev Str.1113 Sofia, Bulgaria; e-mail: detcheva@math.bas.bg*

Major Fields of Scientific Research: Web-based applications, Image processing, analysis and classification, Knowledge representation, Buisiness applications, Applications in Medicine and Biology, Applications in Psychology and Special Education, Computer Algebra.

Mirena Velkova – *Assosiation Autism (President), Sofia, Iskar quarter, Drujhba Komplex, Block 50, Entr. D, App. 101; e-mail: mirena898@gmail.com*

Major Fields of Scientific Research: Autism Treatment, Psychology, Social Studies

Ani Andonova *– Center for Social Rehabilitation and Integration of people with ASD (Director), Sofia, 168 Stamboliisky Blvd. e-mail: ani.andonova@gmail.com*

Major Fields of Scientific Research: Autism Treatment, Psychology, Social Studies, Business informatics, Software technologies

RECURRENT PROCEDURE IN SOLVING THE GROUPING INFORMATION PROBLEM IN APPLIED MATHEMATICS

V. Donchenko, Yu. Krivonos, Yu. Krak

***Abstract**: Number of The grouping information problem in its two basic manifestations recovering function, represented by empirical data (observations) and problem of classification (clusterization) and conception of its solving by the standard recurrent procedures are proposed and discussed. It is turn out that in both case correspond procedures can be designed on the base of so called neurofunctional transformations (NfT—transformations). Conception of such transformations implements the idea of superposition of standard functions by certain sequence of recurrent applications of the superposition. Least Square Method is used for designing the elementary functional transformations and implemented by necessary developed of M-P-inverse technique. It is turn out that the same approach may be designed and implemented for solving the classification problem. Besides, the special classes of beam dynamics with delay were introduced and investigated to get classical results regarding gradients. These results were applied to optimize the NfT—transformations.*

***Keywords**: Grouping information problem, generalized artificial neuronets, learning samples, beam dynamics, Fuzzy likelihood equation, Multiset theory.*

***ACM Classification Keywords**: G.2.m. Discrete mathematics: miscellaneous,G.2.1 Combinatorics. G.3 Probability and statistics, G.1.6. Numerical analysis I.5.1.Pattern Recognition H.1.m. Models and Principles: miscellaneous:*

Introduction

The problem of grouping the information (grouping problem) is the fundamental problem of applied investigations. It appears in various forms and manifestations. All of them eventually are reduced to two forms. Namely, these are: the problem of recovering the function represented by their observations and the problem of clustering, classification and pattern recognition. State of art in the field is represented perfectly in[Kohonen,2001], [Vapnik, 1998], [Haykin, 2001], [Friedman, Kandel,2000], [Berry,2004] .

It's opportune to mark what the information regarding the object or a collection of similar object is exposed to aggregating is. It is of principal importance that an object is considered as a set of its main components and fundamental for the object ties between them. Such consideration and only this one enable application of the math in object description, namely, for math modeling. It is due the fact that after Georg Cantor the objects of investigation in math (math structures) are the sets plus "ties" between its elements. There are only four (may be, five) fundamental mathematical means to describe these "ties". Namely, these are: relations, operations, functions and collections of subsets (or combinations of mentioned above). Thus, the mathematical description of the object (mathematical modeling) cannot be anything other than representing the object structure by the means of mathematical structuring. It is applicable to the full extent to that objects which indicated by the term “complex system”. A “complex system” should be understanding and, correspondingly, determined, as an objects with complex structure (complex “ties”). Namely, when reading attentively manuals by the theme (see, for example, [Yeates, Wakefield, 2004], [Forster , Hölzl, 2004]) one could find correspondent allusions. It is reasonable understanding of “complex systems” instead of the its understanding as the “objects, consisting of numerous parts, functioning as an organic whole”.

So, math modeling is designing in math “parts plus ties”, which reproduce “part plus ties” in reality.

So it is principal question in math modeling which math objects represents “part” of the object and which the “ties” ones. The math object - representative should be chosen in such a way that variety of math structuring means were sufficient to convey the object structure.

It is commonly used approach for designing objects - representative to construct them as an finite ordered collection of characteristics: quantitative (numerical) or qualitative (non numerical). Such ordered collection of characteristics is determined by term cortege in math. Cortege is called vector when its components are numerical. In the function recovering problem objects - representatives are vectors and functions are used as a rule to design correspond mathematical “ties”. In clustering and classification problem the collection may be both qualitative and quantitative. In last case correspond collection is called feature vector. It is reasonable to note that term “vector” means more, than simply ordered numerical collection. It means that curtain standard math “ties” are applicable to them. These “ties” are adjectives of the math structure called Euclidean space denoted be R^n. Namely these are: linear operations (addition and scalar multiplying), scalar product and correspond norm.

Just the belonging to the base math structure (Euclidean space) determines advantages of the “vectors” against “corteges”. It is noteworthy to say, that this variant of Euclidean space is not unique: the space $R^{m\times n}$ of all matrixes of a fixed dimension $m\times n$ may represent alternative example. The choice of the R^n space as “environmental” structure is determined by perfect technique developed for manipulation with vectors. These include classical matrix methods and classical linear algebra methods. SVD-technique and methods of Generalized or Pseudo Inverse according Moore – Penrose are comparatively new elements of linear matrix algebra technique [Nashed, 1978](see, also, [Albert,1972] , [Ben-Israel, Greville ,2002]). Outstanding impacts and achievements in this area are due to N.F Kirichenko (especially, [Кириченко, 1997] [Kirichenko, 1997], see also [Кириченко, Лепеха, 2002]). Greville’s formulas:forward and inverse -for pseudo inverse matrixes, formulas of analytical representation for disturbances of pseudo inverse, - are among them. Additional results in the theme as to furter development of the technique and correspondent applications one can find in [Кириченко, Лепеха, 2001], [Donchenko , Kirichenko,Serbaev, 2004], [Кириченко, Крак, Полищук,2004] [Kirichenko, Donchenko, Serbaev,2005], [Кириченко, Донченко,2005] [Donchenko, Kirichenko , Krivonos, 2007], [Кириченко, Донченко,2007] ,[Кириченко, Кривонос, Лепеха 2007], [Кириченко, Донченко, Кривонос, Крак, Куляс, 2009].

As to technique designing for the Euclidean space $R^{m\times n}$ as “environmental” one see, for example [Донченко, 2011]. Speech recognition with the spectrograms as the representative and the images in the problem of image recognition are the natural application area for the correspond technique.

As to the choice of the collection (design of cortege or vector) it is necessary to note, that good “feature” selection (components for feature vector or cortege or an arguments for correspond functions) determines largely the efficiency of the problem solution.

As noted above, the efficiency of problem solving group, the choice of representatives of right: space arguments or values of functions and suitable families past or range of convenient features vectors. This phase in solving the grouping information problem must be a special step of the correspondent algorithm. Experience showed the effectiveness of recurrent procedures in passing through selection features step. For correspond examples see, [Ivachnenko,1969] with Ivachnenko's GMDH (Group Method Data Handling), [Vapnik, 1998] with Vapnik's Support Vector Machine. Further development of the recurrent technique one may find in Donchenko, Kirichenko,Serbaev, 2004], [Кириченко, Крак, Полищук,2004] [Kirichenko, Donchenko , Serbaev,2005], [Кириченко, Донченко,2005] [Donchenko, Kirichenko , Krivonos, 2007], [Кириченко, Донченко,2007] ,[Кириченко, Кривонос, Лепеха 2007]. The idea of nonlinear recursive regressive transformations (generalized neuron nets or neurofunctional transformations) due to Professor N.F Kirichenko is represented in the works referred earlier in its development. Correspondent technique has been designed in this works separately for each of two its basic form f the grouping information problem. The united form of the grouping problem solution is represented here in further consideration. The fundamental basis of the recursive neurofunctional technique include the development of pseudo inverse theory in the publications mentioned earlier first of all due to Professor N.F. Kirichenko and his disciples.

The essence of the idea mentioned above is thorough choice of the primary collection and changing it if necessary by standard recursive procedure. Each step of the procedure include detecting of insignificant components, excluding or purposeful its changing, control of efficiency of changes has been made. Correspondingly, the means for implementing the correspondent operations of the step must be designed. Methods of neurofunctional transformation (NfT) (generalized neural nets, nonlinear recursive regressive transformation: [Donchenko , Kirichenko,Serbaev, 2004] [Кириченко, Крак, Полищук, 2004], [Кириченко, Донченко, Сербаєв, 2005]).

Neurofunctional transformation in recovering function problem

The fundament of the Math truth is the conception of deducibility. It means that the status of truth (proved statement) has the statement which is terminal in the specially constructed sequence of statements, which called its proof. The peculiarity in sequence constructing means, that a next one in it produced by previous by special admissible rules (deduction rules) from initial admissible statements (axioms and premises of a theorem). As a rule, corresponded admissible statements have the form of equations with the formulas in both its sides. So, each next statement in the sequence-proof of the terminal statement is produced by previous member of sequence (equation) by changing some part of formulas in left or right it side on another: from another side of equations-axioms or equations premises. The specification of the restrictions on admissible statements and the deduction rules are the object of math logic.

As it was already marked, the idea of neurofunctional transformation (NfT-) or neurofunctional transformation in recovering function problem in the variant of inverse recursion was offered in [Кириченко, Крак, Полищук, 2004], and in variant of forward recursion - in [Donchenko , Kirichenko,Serbaev, 2004] , [Кириченко, Донченко, Сербаєв, 2005]. References on neuronets is determined by the fact that NfT generalizes artificial neuronets: in possibilities of the standard functional elements (ERRT(elementary recursive regression transformation) in NfT): in topology of its connection; in adaptive design of NfT structure in the whole; in adequate math for its description. Just this forward variant will considered below. Namely , NfT- is the transformation built by recursive application of the certain standard element, which will be designated by abbreviation ERRT (Elementary Recursive

Regression Transformer). Process of construction of the NfT- transformation consists in connection of the next ERRT (or certain number of it) to already constructed during previous steps transformer according to one of three possible types of connection (connection topology). Types of connection which will be designated as “parinput”, “paroutput” and “seq”, realize natural variants of use of an input signal: parallel or sequential over input, - and parallel over output. An input of the Output of current step of recursion is input of the next step.

The basic structural element of the NfT- -transformer is ERRT - an element [Кириченко, Донченко, Сербаєв, 2005], which is determined as mapping from R^{n-1} in R^m of a kind:

$$y = A_{+}\Psi_u\left(C\begin{pmatrix} x \\ 1 \end{pmatrix}\right) \tag{1}$$

which approximates the dependence represented by training sample

$$(x_1^{(0)}, y_1^0), \ldots, (x_M^{(0)}, y_M^{(0)}),\ x_i^{(0)} \in R^{n-1},\ y_i^{(0)} \in R^m,\ i = \overline{1,M},$$

where:

- C–(n×n) - matrix, which performs affine transformation of the vector $x \in R^{n-1}$ - an input of the system; it is considered to be given at the stage of synthesis of ERRT;
- Ψ_u – nonlinear mapping from R^n in R^n, which consists in component-wise application of scalar functions of scalar argument $u_i \in \Im$, $i = \overline{1,n}$ from the given final set $\Im$ of allowable transformations, including identical transformation: must be selected to minimize residual between input and output on training sample during synthesis of ERRT;
- A+– solution A with minimal trace norm of the matrix equation

$$AX_{\Psi_u C} = Y, \tag{2}$$

in which matrix $X_{\Psi_u C}$ formed from vector-columns $\Psi_u(C\begin{pmatrix} x_i^{(0)} \\ 1 \end{pmatrix}) = \Psi_u(z_i^{(0)})$, and Y – from columns, $y_i^{(0)}, i = \overline{1,M}$.

In effect, ERRT represents empirical regression for linear regression y on $\Psi_u\left(C\begin{pmatrix} x \\ 1 \end{pmatrix}\right)$), constructed with method of the least squares, with previous affine transformation of system of coordinates for vector regressor x and following nonlinear transformation of each received coordinate separately.

Remark 1. Further we shall assume that functions of component-wise transformations from $\Im$ would have a necessary degree of smoothness where it is necessary.

Task of synthesis of ERRT by an optimal selection of nonlinear transformations of coordinates on the given training sample was introduced and solved in already quoted above work [Кириченко, Донченко, Сербаев, 2005]. The solution of a task of synthesis is based on methods of the analysis and synthesis of the pseudoinverse matrices, developed in [Кириченко, 1997]. Particularly, reversion of Grevil’s formula [10] was proved in these works, that recurrently allows to recalculate pseudoinverse matrices when a column or a row of the matrix changed by another one.

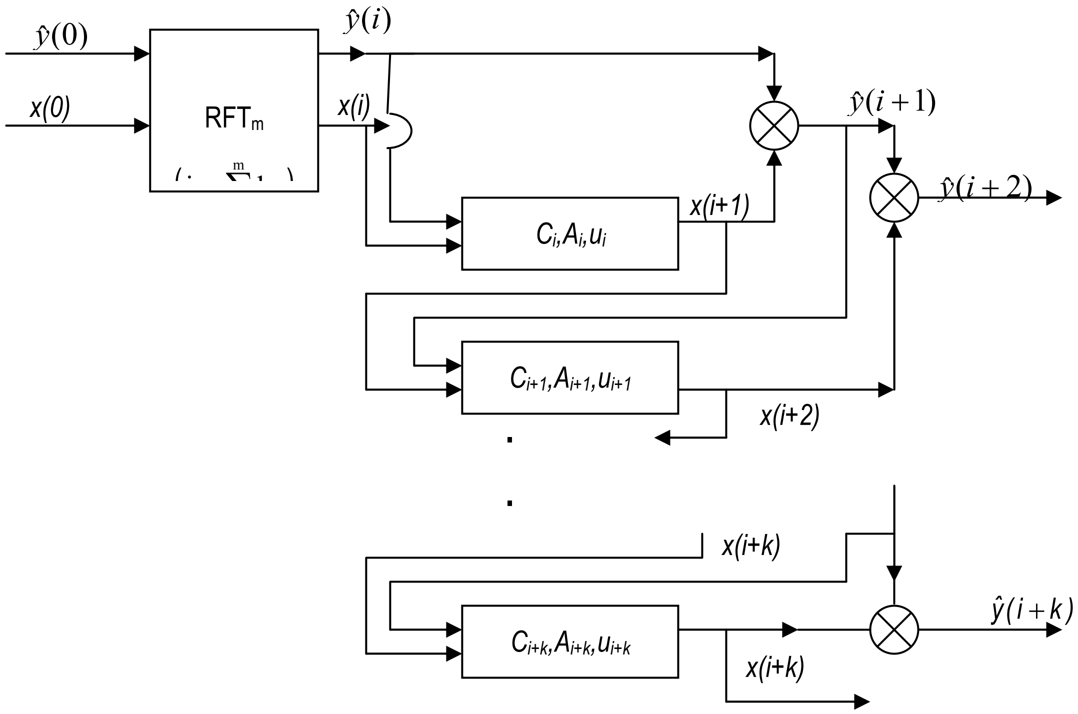
Recurrent procedure in NfT- design: topology and mathematics

Recursion in construction of the NfT--transformer in variant of forward recursion offered below will be considered in the generalized variant in which several ERRTs is used in recurrent connection to already available NfT--structure. Total quantity of recurrent references we shall designate through N, and quantity of ERRTs used on a step m – by $k_m, m = \overline{1,N}$. The common number of ERRTs, used for construction of whole transformer will be designated by $T : T = \sum_{m=1}^{N} k_m$.

Variants of the generalized forward recursion which depend from type of connection of attached ERRT: parinput, paroutput and seq, - are represented on figures 1-3. Where $\hat{y}$ designates an output of already available NfT--structure or an output of the same structure after connection of the next ERRT from the total number k_m of such elements, attached on a step m of the recursion: $m = \overline{1,N}$. Each figure is accompanied by the system of equations which determine transformation of a signal on the next step of recursion.

Type parinput:

Figure 1. Scheme of connection of parinput type – forward recursion.



In parinput type of connection already constructed structure approximates an output of training sample by its input, and set of ERRTs - resulting residual of such approximation depending from an input of training sample. Transformation of the information at this type of connection is described by the following system:

$$\begin{aligned} &x(i+j) = A_{i+j-1}\Psi_{u_{i+j-1}}(C_{i+j-1}x(i)), \\ &\hat{y}(i+j) = \hat{y}(i+j-1) + A_{i+j-1}\Psi_{u_{i+j-1}}(C_{i+j-1}\cdot x(i)) \\ &i = \sum_{l=1}^{m} k_l ,\ j = \overline{1,k_{m+1}} \end{aligned} \quad (3)$$

Type paroutput:

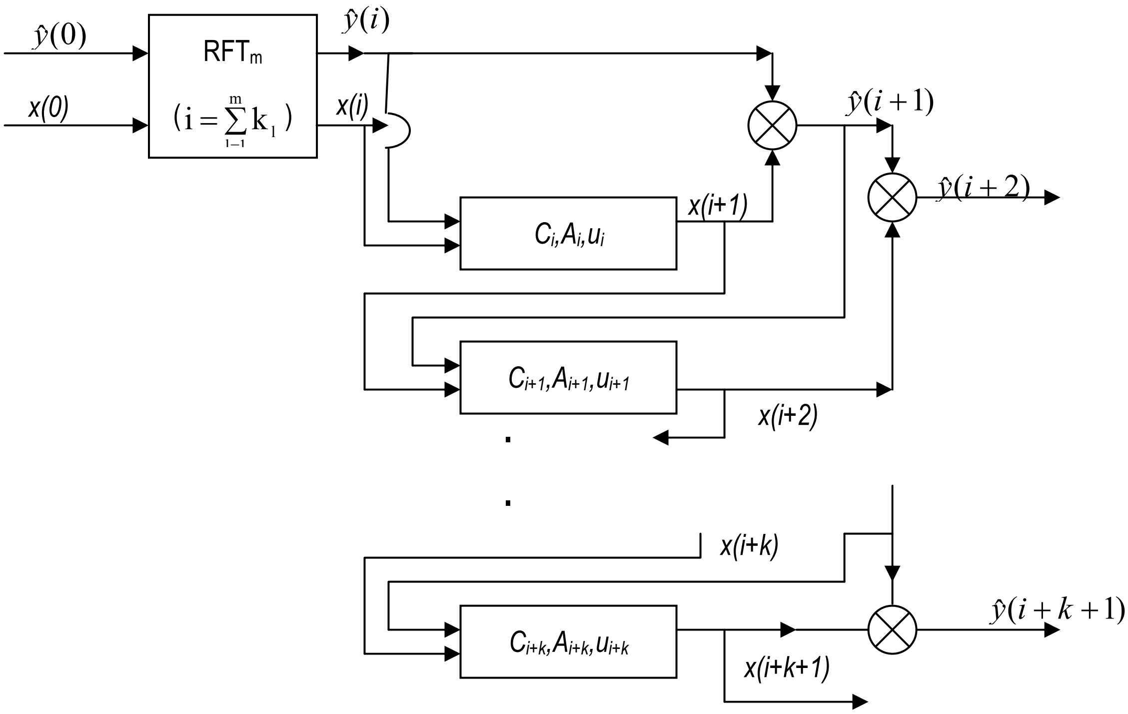


Figure 2. Scheme of connection of paroutput type – forward recursion.

The system describing transformation inside the transformer and communication between an input and an output, at this type of connection looks like:

$$\begin{aligned} &x(i+j)=A_{i+j-1}\Psi_{u_{i+j-1}}(C_{i+j-1}x(i+j-1)),\\ &\hat{y}(i+j)=\hat{y}(i+j-1)+A_{i+j-1}\Psi_{u_{i+j-1}}(C_{i+j-1}\cdot x(i+j-1))\\ &i=\sum_{l=1}^{m}k_l,\ j=\overline{1,k_{m+1}} \end{aligned} \tag{4}$$

Type seq:

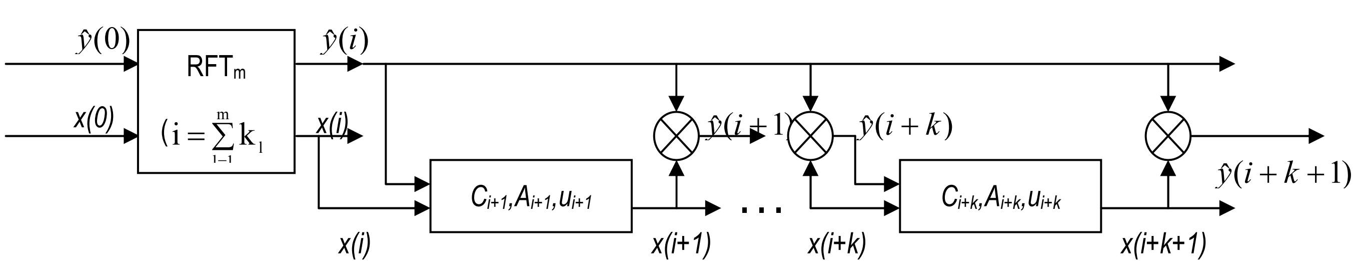


Figure 3. Scheme of connection of seq type – forward recursion.

The equations describing transformation of the information on the next step of recursion look as follows:

$$x(i+j) = A_{i+j-1}\Psi_{u_{i+j-1}}(C_{i+j-1}x(i+j-1)),$$
$$\hat{y}(i+j) = \hat{y}(i) + A_{i+j-1}\Psi_{u_{i+j-1}}(C_{i+j-1}\cdot x(i+j-1)) \quad (5)$$
$$i = \sum_{l=1}^{m} k_l,\ j = \overline{1,k_{m+1}}$$

In this scheme of connection RFT_{m-1} approximates an output part of training sample by input part, and set of ERRTs approximates residual which depends from an output of the next ERRTs.

Entry conditions for all types of connections are described by equations:

$x(0) = x$ – an input of whole NfT--transformer, (6)

$\hat{y}(0) = 0,\ \hat{y}(1) = x(1)$ for all types of connections. (7)

According to (6) in a training mode the inputs of NfT--transformer are $x_{i1}^{(0)} : x_i^{(0)} \in R^{n-1}, y_i^{(0)} \in R^m, i = \overline{1,M}$ and outputs are – $y_i^{(0)} \in R^m, i = \overline{1,M}$.

Connections of recursive construction of the NFT--transformer are determined so, that standard functional of the least squares method is minimized during its construction, i.e.

$$\sum_{i=1}^{M} \| y_i^{(0)} - RFT(x_i^{(0)}) \|^2 \quad (8)$$

Equations (3) - (7) for N steps of recursion with common number T of used ERRTs, $T = \sum_{m=1}^{N} k_m, k_m$,– number of ERRTs, used on a step with number $m = \overline{1,N}$, and also efficiency functional (8) - represent mathematical model of NfT- transformation.

Math for NfT: Discrete control systems with delay

Equations (3)-(8) represent the system of the recurrent equations being certain generalization of a classical control system with discrete time (for details see, for example [Бублик, Кириченко,1975]) and first of all in referring to delay.

The simple and combined control systems with delay

Definition. Simple, accordingly - combined, - nonlinear control system with delay on a time interval $\overline{0,N}$ is a control system whose trajectories are defined by system of recurrent equations (9), accordingly - (10), entry conditions (11) and efficiency functional (12) and represented below:

$$x(j+1) = f(x(j-s(j)), u(j), j) \ , \quad (9)$$

$$x(j+1) = f(x(j), x(j-s(j)), u(j), j) \ . \quad (10)$$

$$j = \overline{0,N-1}, \qquad x(0) = x^{(0)}, \quad (11)$$

$$I(U) = \Phi(x(N)), \quad (12)$$

where function s(j), k=0,...,N–1: s(0)=0, $s(j) \in \{2,...,j\}, j = \overline{0,N-1}, j = \overline{0,N-1}$ – is known.

Evidently, systems with delay for a beam of trajectories are determined. Entry conditions of trajectories of a beam we shall designate $(x(0))_i = x_i^{(0)}, i=\overline{0,M}$, M– number of trajectories of a beam. Trajectories of a beam for both types of systems with delay we shall designate by the appropriate indexation: $x_i(j), j=\overline{0,N-1}, j=\overline{0,M}$,

Let's define efficiency functional for a beam of dynamics which we shall consider dependent only from final states of trajectories of a beam, with equation:]

$$I_p(U)=\sum_{i=1}^{M}\Phi^{(i)}(x_i(N)). \tag{13}$$

Remark 2. Evidently, efficiency functional, as well as for classical control systems, may be defined on all trajectory or trajectories. However, systems with delay at which efficiency functional depends only from final states of a trajectory will be considered in context of NfT--transformers.

Phase trajectories of simple systems with delay are defined by a set of functions f(z, u, j), $j=\overline{0,N-1}$ with one argument z, responsible for a phase variable, and for combined one - a set f(z,v,u,j), $j=\overline{0,N-1}$ with two variables: z, v, which respond for phase variables. Gradients on a phase variables will be denoted by $grad_v f$.

The problem of optimization for both types of such systems with delay is being solved, as well as in a classical case, with construction of the conjugate systems and functions of Hamilton. The assumptions of the smoothness providing correct construction of conjugate systems and functions of Hamilton, and also their use for gradients calculations on controls completely coincide with classical and further will be considered automatically executed.

Optimization in simple control systems with delay

As the analysis of the numerous sources on Theory of Probability and Math Statistics [Донченко 2009], notion of experiment in them is associated with something, named conditions (condition of experiment), under which phenomena is investigated, and something, that appears under the conditions: named the results of experiment.

So, as in [Донченко 2009] "experiment" is proposed to be considered the pair *(c, y)*: *c*- conditions of experiment (observation, trail, test), y – result of experiment. Henceforth Y_c for the fixed condition *c* will denote the set of all possible that may appear in the experiments under conditions $c \in C$. Generally speaking Y_c is not singleton.

It is reasonable to mark out in a condition c variational, controlled, part x: $x \in R^p$ as a rule, and part f, which is invariable by default in a sequence of experiment. Condition c under such approach is denoted be the pair: c=(x, f), $x \in X \subseteq R^p$.

Definition. The conjugate system of a simple control system with delay (9) - (13) we shall call following recurrent equation concerning $p(k): k=\overline{N,0}$:

$$p(k)=\sum_{j\in\{j:j-s(j)=k,\geq kj\}} grad_{x(k)}\{p^T(j+1)f(x(j),u(j),j)\} \quad ,$$

$$k=\overline{N-1,0}$$

with the initial condition

$$p(N)=-grad_{x(N)}\Phi(x(N)).$$

Accordingly, in the case of a beam of trajectories the conjugate systems are defined for each trajectory by equations:

$$p^{(i)}(N) = -grad_{x_i(N)}\Phi^{(i)}(x_i(N)),$$

$$p^{(i)}(k) = \sum_{j\in\{j:j-s(j)=k,j\geq k\}} grad_{x_i(k)}\{p^{(i)T}(j+1)f(x_i(j),u(j),j)\} \quad ,$$

$$k = \overline{N-1,0}, i = \overline{1,M} .$$

Function of Hamilton for simple system with delay is defined by a classical equation:

$$= p^T(k+1)f(x(k-s(k)),u(k),k), k = \overline{N-1,0}$$

For a beam of trajectories of a simple control system with delay the set of functions of Hamilton $\mathrm{H}^{(i)}$, $i = \overline{1,M}$ for each of trajectories $x_i(k), k = \overline{N-1,0}$.

Theorem 1. The gradient on control from the efficiency functional which depends only from final states of trajectories of a beam, for a simple control system with delay is defined by equations:

$$grad_{u(k)}I(U) = -\sum_{i=1}^{M} grad_{u(k)}\left\{p^{(i)T}(k+1)f(x_i(k-s(k)),u(k),k)\right\} =$$

$$= -grad_{u(k)}\sum_{i=1}^{M} H^{(i)}(x_i(k),u(k),k),\ k = \overline{N-1,0}$$

The proof. The proof will be carried out precisely the same as in a classical case: first - for one trajectory, and then by use of additivity of efficiency functional on trajectories of system.

Optimization in combined control systems with delay

Definition . The conjugate system for the combined control system with delay is the system determined by recurrent equations:

$$p(k) = grad_z\{p^T(j+1)f(x(k),x(k-s(k)),u(j),j) \quad +$$

$$+ \sum_{j\in\{j:j-s(j)=k,\geq kj\}} grad_v\{p^T(j+1)f(x(k),x(j),u(j),j)\},$$

$$k = \overline{N-1,0}, i = \overline{1,M} ,$$

with the initial condition

$$p(N) = -grad_{x(N)}\Phi(x(N)) .$$

Accordingly, function of Hamilton *H(p(k+1),x(k),x(k–s(k),u(k),k))* of the combined system is defined by a equation:

H(p(k+1),x(k),x(k–s(k),u(k),k))=

$$= H\left(p(k+1,x(k),x(k-s(k)),u(k),k)\right) = p^T(k+1)f(x(k),x(k-s(k)),u(k),k) .$$

As before, the upper index $(i): i = \overline{1,M}, p^{(i)}(k), H^{(i)}, k = \overline{N-1,0}$, will define objects for trajectories of a beam.

Theorem 2. Gradients on the appropriate controls from the efficiency functional which depends only from final values of trajectories of the combined control system with delay, are defined by gradients from the appropriate functions of Hamilton:

$$grad_{u(k)} I_p(U) = -grad_{u(k)} H(p(k+1), x(k), x(k\text{-}s(k)), u(k)), k))\ .$$

And, hence, for a beam of dynamics the appropriate gradient is defined by equation:

$$grad_{u(k)} I_p(U) = -\sum_{i=1}^{M} grad_{u(k)} H^{(i)}(p^{(i)}(k+1), x^{(i)}(k), x^{(i)}(k\text{-}s(k), u(k), k)) \tag{14}$$

The result may be proved in a standard way for use of the conjugate systems and functions of Hamilton.

NfT and control systems with delays

As it has been already marked NfT--transformation may be represented by a control system with delay. More precisely, the following theorem is valid.

Theorem 3. Regression RFTN-transformer with the direct N-times recursive reference to $k_m, m = \overline{1,N}$ ERRT elements on each of the steps of the recursion is represented by the nonlinear combined beam of dynamics with delay on the set $\overline{0,T} : T = \sum_{m=1}^{N} k_m$:

- with a phase variable $z(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix}$ with $z_i(t) \in R^m, i = 1,2, t = \overline{0,T}$;
- by the system of the recurrent equations which determines trajectories of a beam:

$$z^{(i)}(t+1) = f(z^{(i)}(t), z^{(i)}(t-k(t)), C_t, t) =$$

$$= \begin{pmatrix} f_1(z^{(i)}(t), z^{(i)}(t-k(t)), C_t, t) \\ f_2(z^{(i)}(t), z^{(i)}(t-k(t)), C_t, t) \end{pmatrix}, \qquad)$$

with $f_1, f_2 \in R^m, = \overline{1, T-1}, i = \overline{1,M}$, , dependent on topology of the NFT--transformer,

- initial condition

$$z^{(i)}(0) = \begin{pmatrix} x_i^{(0)} \\ 0 \end{pmatrix}, i = \overline{1,M} ,$$

where $z^{(i)}(0), i = \overline{1,M}$, initial conditions of trajectories of a beam, and $x_i^{(0)}, i = \overline{1,M}$,– elements of an input of learn sample;

- efficiency functional $I(C), C = (C_1, \ldots, C_T)$, which depends on matrices $C_1, \ldots, C_T$ as on controls:

$$I(C) = \sum_{k=1}^{M} \| y_k^{(0)} - z_2(T) \|^2 ,$$

where $y_k^{(0)}, k = \overline{1,M}$, components of an output of learning sample.

The proof can be found in [Кириченко, Донченко, Сербаев 2005].

The statement proved enables using of methods of optimization of the theory of control for optimization of already constructed RFTN-transformer on residual size depending on matrices $C_0, C_1, ..., C_{T-1}$. This statement also is a subject of the following theorem.

Theorem 4. By presence of continuous derivatives up to the second order inclusive of functions of family $\Im$ RFT - transformation may be optimized by gradient methods with gradients of the efficiency functional on matrices C, as on parameters.

Proof. According to the theorem 3 NFT--transformation may be represented by the combined control system with delay, and according to the theorem 2 gradients on the appropriate controls - parameters of the NfT--transformation are defined by equation (14). Importance of the given theorem lies in that it gives exhaustive interpretation of Back Propagation algorithm, outlining at the same time borders of the specified method.

Neurofunctional transformation in classification problem

Clusterization and classification problem as the variant grouping problem will be discussed according [Кириченко, Кривонос, Лепеха, 2007] for two classes $\Omega_x(1), \Omega_x(2) \subseteq R^m$ and, correspondingly, with two correspondent learning samples $x(j) \in \Omega_x(1), j \in J_1$, $x(j) \in \Omega_x(2), j \in J_2 : J_1 \cup J_2 = \{1,...,n\}, J_1 \cap J_2 = \varnothing$. The classification problem will be interpreted as the problem of designing function $\varphi : \varphi \in R^m \to R^1$ (discriminate function), which would "Δ – differentiate" classes for some $\Delta > 0$, in the sense, that:

$\varphi(x(j)) \geq \Delta, j \in J_1, \varphi(x(j)) \leq -\Delta, j \in J_2$

We will find correspond φ, Δ for linear case: when $\varphi(x) = a^T x, a = (a_1, ..., , a_m)^T \in R^m$ (Linear Discrimination Problem (LD -problem)):

$$a^T x(j) = y(j) \geq \Delta, j \in J_1 \quad (15)$$

$$a^T x(j) = y(j) \leq -\Delta, j \in J_2 : \quad (16)$$

$$J_1 \cup J_2 = \{1, ..., n\},$$

$$J_1 \cap J_2 = \varnothing .$$

Under matrix denotation:

$$X = \left(x(1) \vdots \ldots \vdots x(n) \right) = \begin{pmatrix} x_{(1)}^T \\ \ldots \\ x_{(m)}^T \end{pmatrix}, x(j) \in R^m, j = \overline{1, n}, x_{(i)} \in R^n, i = \overline{1, m},$$

$$y = (y(1), ..., y(n)^T \in R^n$$

LD-problem (15),(16) one can rewrite in matrix form under

$$X^T a = y : y \in \Omega_y(\Delta) \subseteq R^n, \Delta > 0, \quad (17)$$

$$\Omega_y(\Delta) = \{ y \in R^n : y_j > \Delta, j \in J_1, y_j < \Delta, j \in J_2 \} \quad (18)$$

Due to (17), (18) LD-problem we can consider as the solving problem for a System of Linear Algebraic Equaitions (SLAE) with plural form of constraint on $y : y \in \Omega_y(\Delta)$.

Due to results of [Kirichenko, 1997], [Кириченко, Лепехаб2002], see also [Кириченко, Донченко,2005], solvability condition for SLAE (17) is of the next form:

$y^T Z(X) y = 0, Z(X) = I_n - X^+ X$

X^+ - Moore-Penrose pseudoinverse (MP-inverse)(see, for example ,[Albert, 1972]).

Thus, the next lemma is valid.

Lemma 1. Solvability SLAE (17) with plural constraint (18) is equivalent to solvability problem for quadratic equation with constraint of the next form:

$$y^T Z(X) y = 0, \Omega_y(\Delta) \subseteq R^n : \Delta > 0 . \quad (19)$$

Due to results of the publication cited before previous Lemma, we have that next statement is valid.

Lemma 2. For each a solution y_d of (19) LD-problem solution is determined by relation

$$a = X^{+T} y_d \quad (20)$$

Thus the next theorem is valid due Lemmas 1,2.

Theorem 5. LD-problem (17)-(18) equivalent to solvability quadratic optimization problem for minimization of quadratic form $y^T Z(X) y$ in domain $\Omega_y(\Delta)$:

$$y_* = \arg \min_{y \in \Omega_y(\Delta)} y^T Z(X) y . \quad (21)$$

Insolvability of the optimization problem from Theorem 5 means insolvability LD-problem with the feature vector of the model. So the features need purposeful change. So, criteria for the choice of correspondent components and means for correspondent changes must be available. Just these means may be realized by the correspondent modification of NfT.

Criteria of informative content for the components of feature vector

There are three criteria for detection of informative value (I-value) for the components of feature vector.

By the first of them component with minimal I-value is the one correspond the number, determined by the relation

$$i^* = \arg \min_{i=\overline{1,m}} y_*^T Z(X_i) y_* . \quad (22)$$

Another criterion is the one determined by the degree of independence of the component from all others:

Coordinate *i* is I-valuable, if

$$x_{(i)}^T q(i) = \begin{bmatrix} 1, \text{feature is I -valuable} \\ \neq 1, \text{feature is I -valuable} \end{bmatrix}, i = \overline{1,m}, \quad (23)$$

where

$X^+ = (q(1),...,q(m))$.

This criterion based on M-P inverse where (22) (see,[Kirichenko,1997]) is criterion of linear independence of *i*-component from the rest.

The next criterion is the one for detection minimal I-value component from dependent component of feature vector. The number *i** of the correspondent feature vector component is determined by relation (see,[Kirichenko,1997]):

$$i^* = \arg\min_{i=\overline{1,m}} \frac{\left| y_*^T q(i) \right|^2}{\left\| q(i) \right\|^2} . \qquad (24)$$

Algorithm of modification for components of feature vector with minimum of informative content

If the component with minimum informative content is detected it must be changed by its modification Other components of feature vector may be used for such modification. Nonlinear transformation ψ of the component by itself or other components may be used for the modification from certain collection, for example from Table 1:

Table 1.Basic nonlinear transformation

ψ	ψ	ψ
$y = \frac{1}{ax+b}$	$y = \frac{a}{x} + b$	$y = \frac{x}{ax+b}$
$y = \frac{1}{ax^2+bx+c}$	$y = \frac{x}{ax^2+bx+c}$	$y = a + \frac{b}{x} + \frac{c}{x^2}$
$y = ax^b$	$y = ab^x$	$y = ae^{bx}$
$y = ae^{-bx^2}$	$y = ax^b e^{cx}$	$y = ae^{bx+cx^2}$
$y = a\sin(bx+c)$	$y = th(ax)$	$y = Arth(ax)$

Algorithms for modification consists of the next steps.

1. Cycle of solution for task (21).
2. Detection the component, say x_s , for modification according to one of the relations (22)-(24).
3. Modification of the detected component standard procedure, which include changing that component by another: for example $\psi(x_s)$, followed by the choice of best $\psi^* \in \Psi$ according to solution of optimization task:

   $\psi^* = \min_{\psi \in \Psi} y_*^T Z(X_{(s,\psi(x_s))}) y_*$,

   where $X_{(s,\psi(x_s))}$ - matrix, which corresponds to new feature vector with feature $\psi(x_s)$ instead of x_s .
4. New cycle of solution for task (21).

Scheme of the algorithm is represented on Figure 1:

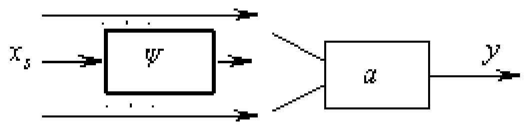


Figure 1.

Modification of step 3 may be of more complicated: with using others of components. For example, next chart (Figure 2) depicts using of nonlinear transformation of another component for changing x_s:

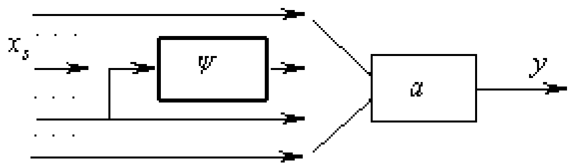


Figure 2.

Some others variants of the modification are represented on the next charts

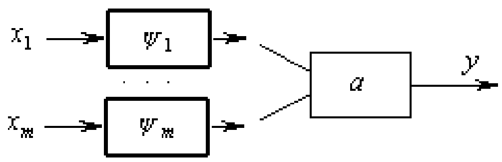


Figure 3.

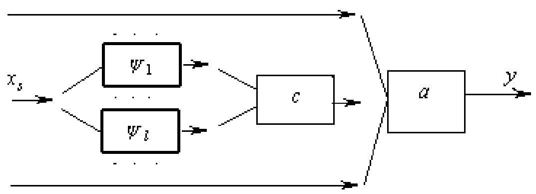


Figure 4.

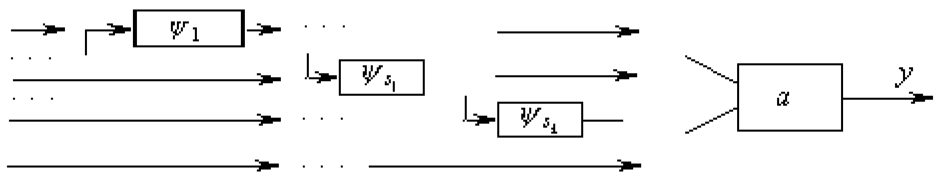


Figure 5.

Chart from Figure 3 represent simultaneous nonlinear transformations for the several component, chart from Figure 4 – modification by changing the component by linear combinations of nonlinear transformations some of others components of feature vector, and Figure 5 represents consequent application of some of one step modification. There may be others variants from those, one can find in [Кириченко, Кривонос, Лепеха, 2007].

Namely charts depicted earlier illustrate the idea of NfT for the transformations of the feature vector.

Conclusion

The realization of the conception for recurrent procedures in solving of important applied grouping information problems is represented. The approach proposed in the article is developed for the both basic form of grouping problem. The development of M-P inverse technique due to Professor Kirichenko and his disciples is the basis for all results of the article.

Bibliography

[Albert ,1972] Albert A.E. Regression and the Moore-Penrose pseudoinverse.-Academic Press.- 1972 .- 180 стор.

[Ben-Israel , Greville ,2002] Adi Ben-Israel ,Thomas N.E. Greville Generalized Inverses .Theory and Applications.- Second Edition.- Springer-Verlag New York, Inc. -2003.-420 p

[Berry,2004] Michael W. Berry Editor Survey of Text Mining:Clustering, Classification, and Retrieval Springer-Verlag New York, Inc. -2004.-244 p.

[Donchenko , Kirichenko,Serbaev, 2004] Donchenko V.S. Kirichenko M.F., Serbaev D.P. Recursive regression transformation and dynamical systems// Proceedings: of the Seventh International Conference “Computer Data analysis and Modeling: robustness and computer intensive methods”. – V.1. – September 6-10, 2004. – Minsk. – P. 147-151

[Donchenko, Kirichenko , Krivonos, 2007] [Donchenko V. Kirichenko M., Krivonos Yu. Generalizing of neural nets: Functional nets of special type// International Journal Information Theories and Applications. – N 3. –V.14. – 2007. – P. 259-266.

[Кириченко, Донченко,2007] Кириченко Н.Ф., Донченко В.С. Псевдообращение в задачах кластеризации.- Кибернетика и системный анализ.- 2007.- № 4.- с. 73 –92

[Friedman, Kandel,2000] Menahem Friedman Abraham Kandel Introduction to pattern recognition statistical, structural, neural and Fuzzy logic approaches world scientific publishing co. Pte. Ltd..-reprinted 2000.-329 p.

[Forster , Hölzl, 2004] John Forster and Verner Hölzl Applied Evolutionary economics and complex systems.-2004-- 293 p.

[Haykin, 2001] Haykin S. Neural Networks, A Comprehensive Foundation.-Prentice Hall, Upper Saddle River, New Jersey 07458. – 1999.– 842 p

[Kohonen,2001] Kohonen T. Self-Organizing Maps -3-d ed. - Tokyo: Springer, 2001.-501 p

[Kirichenko, Donchenko , Serbaev,2005] Kirichenko N.F., Donchenko V.S. Serbaev D.P. Nonlinear recursive nonlinear Transfomations: Dynamic sestems and Optimizations.//Cybernetics and Systen Ansalysis.– V.41, №3.– May 2005.– p.364-373.

[Kirichenko, 1997] N.F. Kirichenko. Analytical Representation of Perturbation of Pseudoinverse Matrices. Cybernetics and Systems Analysis. Vol.33, Number 2. P.230-239. March-April 1997.

[Nashed,1976] Nashed M. Zuhair ,Votruba G.F., editors. A Unified Operator Theory of Generalized Inverse.//Proceedings of an Advanced Seminar Sponsored by the Mathematical Research Center, The University of Wisconsin, Madison, October 8-10, 1973. – New York, Academic Press, 1976.–1200 p.

[Vapnik, 1998] Vapnik, V.N. Statistical Learning Theory. New York: Wiley, 1998

[Yeates, Wakefield, 2004] Donald Yeates, Tony Wakefield Systems Analysis And Design . -

Pearson Education. – 2004. - 499 p.

[Ivachnenko, 1969] Івахненко А.Г. Системи розпізнавання і автоматичного управління, що Самоорганізующиеся. – К.: Техніка, 1969.–395 с.

[Бублик, Кириченко,1975] Бублик Б.Н., Кириченко Н.Ф. Основы теории управления. – К.: Высшая школа, 1975.–328 с.

[Донченко,2011] Владимир Донченко Евклидовы пространства числовых векторов и матриц: конструткивные методы описания базовых структур и их использование.// International Journal “Information technologies & Knowledge”.- 2011.- Vol. 5.- Number 3.-P.203-216.

[Кириченко,1997] Н.Ф. Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц.// Кибернетика и системный анализ. – 1997. – № 2. – С. 98-107.

[Кириченко, Лепеха, 2002] Кириченко Н.Ф., Лепеха Н.П. Применение псевдообратных и проекционных матриц к исследованию задач управления, наблюдения и идентификации.// Кибернетика и системный анализ. – 2002. – № 4. – С. 107-124.

[Кириченко, Крак, Полищук, 2004] Кириченко Н. Ф., Крак Ю. В., Полищук А.А. Псевдообратные и проекционные матрицы в задачах синтеза функциональных преобразователей.// Кибернетика и системный анализ –2004.–№3.

[Кириченко , Донченко, Сербаев, 2005] Кириченко Н.Ф., Донченко В.С., Сербаев Д.П. Нелинейные рекурсивные регрессионные преобразователи: динамические системы и оптимизация.// Кибернетика и системный анализ.– №3, 2005.– С. 58-68.

[Кириченко, Кривонос, Лепеха 2007] Кириченко Н.Ф., Кривонос Ю.Г., Лепеха Н.П. Синтез систем нейрофункциональных преобразователей в решении задач классификации. // Кибернетика и системный анализ. – 2007. – №3. С. 47-57

[Кириченко, Донченко, Кривонос, Крак, Куляс, 2009] Кириченко М.Ф., Донченко В.С Кривонос Ю.Г., , Крак Ю.В.,., Куляс А.І. Аналіз та синтез ситуацій в системах прийняття рішень.- Київ: Наукова думка.-2009. 336с.

[Кириченко, Донченко,2005] Кириченко М.Ф., Донченко В.С. Задача термінального спостереження динамічної системи: множинність розв’язків та оптимізація//Журнал обчислювальної та прикладної математики. – 2005. –№5– С.63-78.

Authors' Information

Volodymyr Donchenko *– Professor, National Taras Shevchenko University of Kyiv. Volodymyrs’ka street, Kyiv, 03680, Ukraine; e-mail: voldon@bigmir.net.*

Yuri Krak– *Professor, National Taras Shevchenko University of Kyiv.*

Yuri Krivinos - *Deputy Director, Institute of cybernetics of NASU; Professor; academician, NASU*

CHOICE OF DIAGNOSTIC DECISION MAKING IN MEDICINE AND INTERVENTION MISTAKE PREDICTION USING MATHEMATICAL MODELS

Ivan Melnyk, Rostyslav Bubnov

***Abstract**: Most processes, found in medicine, are nonlinear, chaotic, have a high level of complexity. The decisions in health care are often stereotyped, managed by habits preferences, previous experience and official directives. These decisions might be not completely conscious. There are a lot of papers, devoted to modeling diagnostics or treatment conduction, but still behavior responses of medical practitioners were not studied, no universal comprehensive and effective model was created. Besides nonlinear nature of biomedical phenomena, pathologies, its chaotic expression, all the information process in medicine at each of its stages, including information perception by available diagnostic tools, analysis, decision making and implementation of therapeutic interventions, are complex, chaotic. We made attempts to integrate this process, bringing scheme into harmony. Each stage requires creation some mathematical model, that might be described by generalized equation. These equations can be substituted into one, that could be solved in closed system. We do not aim to find some absolute kind of decision, its statistically calculated optimal way of solution, but accent on a special mood, the state of expert, which could give a possibility to make only one correct decision with failure in input parameters. In such cases the lack of prior data is compensated by doctor's experience.*

***Keywords**: imaging, mathematical modeling, intervention, choise, error analysis, Monty Hall paradox, method of branches and boundaries.*

***ACM Classification Keywords**: H. Information Systems: H.1 MODELS AND PRINCIPLES: H.1.0 General; G.1.0 Mathematics of Computing General Error analysis; G.2 DISCRETE MATHEMATICS: G.2.1 Combinatorics: Combinatorial algorithms; G.2.2 Graph Theory.*

Introduction

The decisions in health care are often stereotyped, managed by habits preferences, previous experience and official directives. Creating a reliable mathematical models and use of information technology at all stages of the treatment process from the expression the pathological processes to implementation of therapeutic interventions associated with neuro-physiological perception of these phenomena, making decisions in the absence of input parameters for creation self-controlled systems based on forecasts of future medical errors are important tasks [1]. Although mathematical models cannot replace human judgment in the field of medicine, they can be useful and crucial to make only one correct decision in the absence of information on input parameters. Until now its often compensated by doctor's experience.

A. Logistic model selection diagnostic decisions in medicine. In the clinical setting, often there are situations, when there is a need for a solution that lies in the choice between several equally probable options. This medical decisions, based on data from scientific studies that have presented statistical calculations, e.g., accuracy, specificity, predicative importance of diagnostic methods and effectiveness of treatment. For lack of input parameters previous experience of the doctor is used and an intuitive decision is made. Often, after the selection process for diagnosis or treatment there is additional information that does not directly affect the preselection process. Such information is often ignored by doctors, is not used to correct medical decisions. To ensure and optimize logically and intellectually controlled diagnostic process we propose scheme for appropriate choice of diagnostic decisions.

We suggest the use of logistic models of Monty Hall paradox and its generalization (in the case of four factors) [2, 3] for optimizing diagnostic decisions in medicine [4]. As certainly, Monty Hall paradox (for three factors) in the primary (classical) formulation taught by the example of three doors. We formulate the statement in more usual language. There are three boxes. In one of the boxes are expensive valuables, and in the other two - less. Two people take part in the performance: the person who chooses the casket (participant), and the person conducting the procedure of choice (presenter). Participant selects one of three boxes in the first step. After that, the presenter chooses among the two remaining boxes, which have smaller values and opens it. He offers to change participant's choice and select new box that has not yet been elected. The question arises: *do not increase the probability to choose casket of precious values, if the participant choosing the proposal will lead?* Thus, **YES** it will increase and it will be as high as **2/3**. If the participant's selection will not be changed, the probability that the casket, which he chose is expensive, has value of **1/3** [2]. This finding contradicts the everyday intuitive perception of most people, so this problem is called *Monty Hall paradox*.

Monty Hall paradox itself is applied to the case of three boxes. There is a practical need to consider the case of four boxes. There are four boxes. In one of the boxes are expensive valuables, and in the other three - smaller values. For this task logistic model selection boxes offer the best values in two stages.

The first stage of the model. Participant selection in the first step chooses one of the four boxes. Three boxes remain that have not yet elected. The presenter chooses (among the three / not yet chosen) box with a smaller value and opens it. Participant is proposed to change the choise and selects the small box with two that have not yet chosen. If he does not change his original choice, the procedure of choice ends. He gets casket, which he chose from the beginning. If he agrees with the change of the initial choice, then - go to the second stage model.
The second stage of the model. Presenter proposes to make the final choice of two boxes that were not elected at the first stage and choose one of them opening it. In this logistic model for the four boxes is an element of paradox. If he does not change his original choice this time, the probability that the casket, which he chose from the beginning, containing winning value is **0.25**. If he changed the initial choice and the second phase selects one of two boxes remaining, the probability of win will be equal to **0.375** [3].

Example (treatment algorithm) [4]. Condition: the treatment started conducting according to one scheme, should not be changed. Several (3 or 4) equivalent circuits according to input parameters are possible. Additional information (such as laboratory tests), which is not directly related to the selected scheme, does not indicate the correct circuit, but can eliminate one or two circuits, while not affecting the choice between those that remained. Change according to the previous selection leads to increase the probability of correct choice from **1/3 to 2/3** (for 3 equivalent schemes), or from **0.25 to 0.375** (equivalent to 4 circuits), for these conditions simulated situation. The features inherent to modern medicine indicate that the appearance of new additional excluding parameters, based on ignorance of the obviously negative option, is often the most randomized. Ignorance of correct choice (additional information regarding all options simultaneously), for example, increases the probability not to **2/3,** but only **to 1/2** for the three schemes.

The study of adverse negative prognostic parameters of interventional mistakes using the scheme of the method of branches and boundaries.

Conducting minimally invasive interventions under radiology / ultrasound control requires continuous improvement of multidisciplinary approach to the analysis of errors and develop a differentiated approach to each clinical situation for achieve the efficiency about 100%. Previously we reported [5,6] to solve combinatorial (correctable) problem of selection options of negative prognostic indicators for interventional radiology / sonography mistakes to ensure a high level of patient safety, as well as study-level skills and minimal training required for training programs for interventional medicine (in particular in pain management) by applying the method of branches and boundaries. From the formal (mathematical) point of view the problem of negative options selection of prognostic indicators for interventional sonography mistakes is a discrete-combinatorial. Finding "good" solutions for such problems usually are resistive in nature. In mathematical terms the problem of finding solutions to these problems are called in the theory of optimal solutions of discrete optimization problems.

Experimental studies. According to the goal we included 2 groups of physicians: 6 anesthesiologists, who had no previous experience in interventional sonography and a group of experts (ultrasound doctors) - 6 people with previous experience in interventions under ultrasound control. Fundamental difference between skills level of ultrasound doctors were excluded, all studies were conducted in relative isolation. Ultrasound examination was carried out using a portable ultrasound device Sonosite M-Turbo with multifrequency linear and convex probes (used in hospital operating room). The study was conducted on special designed phantoms, which included a gel phantom, phantom and biological electronic device to record accurate needle penetration into the object. All professionals - ultrasound doctors (experts) and anesthesiologists (novices) performed 30 punctures of each group of studies. The comparative study of different methods of introducing the needle to different kinds of phantoms was conducted, recording performance, mistakes were determined, statistical analysis was performed. In case of absence of experimental data for the formation of separate branches of the graph the expert method according to clinical experience of two independent experts was applied.

One approach to solving discrete optimization problems are algorithmic scheme of the method of branches and boundaries. For the first time this method was proposed by Land and Doig [7] in 1960 for solving integer linear programming. When applying algorithmic scheme of the method of branches and boundaries to solve a specific class of discrete optimization to use mathematical characteristics and specificity of this class of problems that often allows us to develop efficient numerical algorithms for the special method of branches and boundaries to address these problems. At the core algorithmic scheme of the method of branches and boundaries is the idea of successive breaking the current set of admissible solutions to a subset (a subset of branching). At each step of this method of partitioning elements (ie subsets of solutions) are checking to determine whether this subset contain the optimal solution. Verification carried out by calculating the value of the lower estimates (lower bounds) objective function (for minimization problem) or the upper estimate (estimates down) objective function (for maximization problem) in this subset of solutions and comparing the value assessment of the value of record at the moment. *Record* - is currently the best objective function value of the found solutions.

For the problem of maximizing algorithmic scheme of the method of branches and boundaries will be as follows. If the upper bound of objective function for this subset of solutions is more (less) record, this subset may be rejected for further consideration, since it obviously does not contain an optimal solution (it is not "promising" for further consideration. Record value will change, if the objective function for the new solution found less than previously estimated record, this new found. If at some step can discard all the elements of partition (a subset of solutions), the record value - an optimal solution of the initial value problem . Otherwise, with subsets of solutions that are not rejected, was elected one of the "promising" and it is divided into subsets of branching. These new solutions again tested a subset of "optimality" and so on, until at some step does not work, meaning that a record will be higher (not lower) values of upper bounds of objective function on all subsets of branching. the end of the computation process and record the current value is the optimal value of objective function and the corresponding solution is the optimal solution of original problem.

The method of branches and boundaries includes two components of treatment: the construction of branches and computation limits (upper) values of the function objective optimization. Branching - is to identify all possible options so as not to leave without loss of any option. We're building a tree, all branches (branching).When you start branching in any situation, the detected branches contain all possible ways of development of the situation. The main requirement is that these subsets do not overlap and their union would have created a whole set of options for solutions. If not cut off branches to complete their analysis, the method branches stood to be exhaustive of all options.

The second component of the procedure of the method of branches and boundaries - the definition and use of boundaries (top) values of the function purpose - to assess their branches without detailed analysis and cut-off "unpromising." Must be, at any time of analysis, numerical rating desired objective function value.

In general, the discrete optimization problem is formulated as follows: to find optimum (maximum of) functions, where the element is selected from some discrete set, is considered such a problem:

$$F(x) \rightarrow \min , \quad (1)$$

$$x \in X , \quad (2)$$

where X - is a discrete set.

The study problem (the problem of the choice options of negative prognostic indicators mistakes interventional sonography) is regarded as a discrete optimization problem (in fact, a problem on a graph) and its solution using algorithmic scheme of the method of branches and boundaries, and procedures strings of finding sequences of branches (by arcs) on the graph [5.6].

The task of selecting variants of negative prognostic indicators of error is interventional sonography optimization (maximization) as a functional solution - the probability of its realization. This problem has a discrete nature and relates to the so-called - full of problems. Find the solution of such problems is resistive in nature and is very time-consuming computing process.

The function aims will be to maximize the likelihood of a decision (the way from the beginning of the branching tree for sequences of branches to the top of the latest branches), that is.

MAX

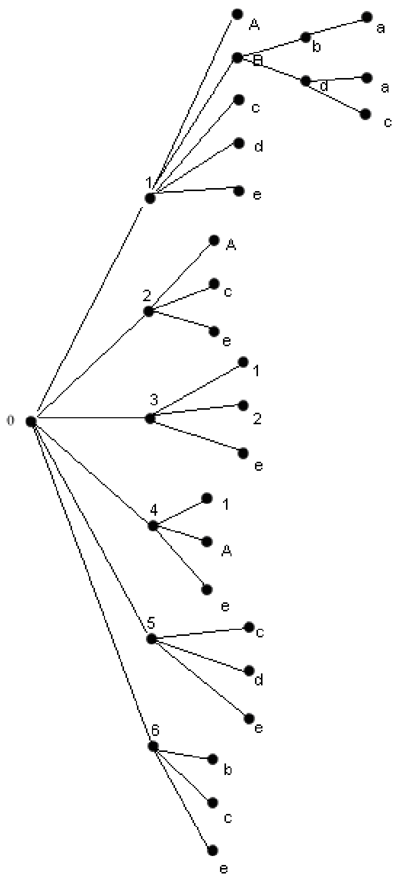
$$F((i_1, i_2),(i_{2,} i_3),...(i_{k-1}, i_k)) = PR(p_{12} x p_{23} x ... x p_{k-1,k}) \quad (3)$$

Where $(i_1, i_2),(i_2, i_3),...,(i_{k-1}, i_k)$ - a sequence of arcs of the tree-graph partial variants (single) decisions are the way to the top Initial (tree roots) to the top of the latest in this way (since it does not have arcs that would come out) - the corresponding probabilities of (appearance) arc path and - the operator multiplying the relevant numbers.

To construct the relevant pathways to solve the problem, (3) are used as mathematical and information technology finding the shortest admissible paths in the graph [8].

Results.

All doctors, who participated in the study, were succeeded in the imaging and intervention trials. The results of the punctures and registered errors of interventions are presented in Fig. 1 (corresponding to part of the graph).



*

Fig. 1. A. Tree (graph) of errors while performing interventions – regional anesthesia under ultrasound guidance (*picture is presented reduced without excessive branching options).*

Thus, on a tree (graph) is shown the algorithm of options (string) negative consequences when performing regional anesthesia with admitted intervention and imaging errors, where

1 - loss of visualization a needle;
2 - loss of visualization an object;
3 - incorrect mapping of testing area and images on the screen;
4 - poor selection of a needling place;
5 - uneven distribution of local anesthetic;
6 - fatigue;
A - incorrect needle trajectory
B - damage to surrounding tissue
a - no effect;
b - reduced quality of anesthesia;
c - longer duration of manipulation;
d - alarm continued correction (eg, third injection).
e - fast adequate correction;

Conclusion

Using these logistic models in clinical practice should optimize the clinical algorithms processing to reduce stereotypical judgments and redundant diagnostic and therapeutic medical procedures. The modeled mistakes led to a decrease in the quality of intervention, but could cause iatrogenic injury in clinical conditions. The method of branches and boundaries effectively solves the problem of choice and interventional mistakes negative prognostic indicators as a discrete optimization problem.

Bibliography

1. [Bubnov, 2011] Bubnov R.V., Melnyk I.M. QUESTION OF CHOISE, DECISION MAKING IN DIAGNOSTIC IMAGING AND INTERVENTION MISTAKE PREDICTION MATHEMATICAL MODELING APPROACH / EPMA Journal (2011) 2 (Suppl 1): S186.
2. [vos Savant,1995], Marilyn vos Savant Review of The World's Most Famous Math Problem "American Mathematical Monthly (The American Mathematical Monthly, Vol. 102, No. 5) 102 (5): 470-473.
3. [Melnyk I.M., 2012] Melnyk I.M. Synthesis logistic model Monty Hall paradox and its use / Materials III All-Ukrainian scientific conference "Information and System Sciences" (Poltava, 1-3 March 2012), Poltava: PUET, 2012. - P. 189-191. (In Ukrainian).
4. [Bubnov, 2012] Bubnov R.V., Melnyk I.M. Application of the logistic model Monty Hall paradox and its generalization for optimizing diagnostic decisions in medicine / Proceedings of the Third All-Ukrainian scientific conference "Information and System Sciences" (Poltava 3.1 March 2012)-Poltava: PUET, 2012. - P. 29-33. (In Ukrainian).
5. [Melnyk, 2011] Melnyk I.M., Bubnov R.V. Solving the problem of studying the adverse prognostic parameter errors regional anesthesia using algorithmic scheme of the method of branches and boundaries / Decision Support Systems. Theory and Practice: Collection of scientific and practical conference with international participation. - Kyiv: IMMSP Nano, 2011. - P. 168-171. (In Ukrainian).
6. [Melnyk, 2011] Melnyk I.M., Bubnov R.V. Application of the method of branches and boundaries of discrete optimization to solve the problem of finding a negative prognostic parameter errors regional anesthesia. Proceedings of the All-Ukrainian scientific seminar "Combinatorial optimization and fuzzy sets" Poltava, 26-27 August 2011. - P. 13-22. (In Ukrainian).
7. [Land, 1960] Land A.H., and Doig A.G. An automatic method of solving discrete programming problems. Econometrics. V. 28 (1960), pp. 497-520.
8. [Yermoliev, 1967] Yermoliev Yu.M., Melnyk I.M. Extreme problems on the graphs. Publishing "Naukova Dumka", K. - 1967. -159 p. (In Russian).
9. [Bubnov, 2011] Bubnov R.V., Strokan A.M. The efficacy study conduct regional anesthesia under ultrasound control on their own phantoms / Proceedings of the Second International Conference "Biomedical engineering and technology" [text]: collection. materials. - K: "KPI", 2011. - P. 23-25. (In Ukrainian).

Authors' Information

Melnyk Ivan M. *, PhD - senior scientist, International Research and Training Center for Information Technology and Systems NAS of Ukraine, Kyiv, Ukraine, 03187, Kyiv, Glushkova ave., 40; e-mail: ivanmelnyk@ukr.net*

Major Fields of Scientific Research: theory for optimization complex systems, genetic algorithm, combinatory and discrete optimization, mathematical modeling for decision making in medicine, economy, fractal analysis.

Bubnov Rostyslav V., *MD, PhD – doctor, Center of ultrasound diagnostics and interventional sonography, clinical hospital "Pheophania" of State Affairs Department, National Representative in Ukraine of the European Association for Predictive, Preventive and Personalised Medicine (EPMA) and Coordinator of the regional EPMA-BOARD in Ukraine; Kyiv, Ukraine; 03680, Kyiv, Zabolotny ave., 21; e-mail: rostbubnov@gmail.com; http://www.rostbubnov.narod.ru/*

Major Fields of Scientific Research: personalized, model guided medicine, radiology, ultrasonography, minimally invasive, interventions, pain medicine, patient safety, image processing, decision making in medicine, fractal analysis.

THE INVERSE METHOD FOR SOLVING ARTIFICIAL INTELLIGENCE PROBLEMS IN THE FRAMEWORKS OF LOGIC-OBJECTIVE APPROACH AND BOUNDS OF ITS NUMBER OF STEPS

Tatiana Kosovskaya, Nina Petukhova

***Abstract**: The paper is devoted to the modifying of Maslov inverse method for a special form of predicate formulas used in the solution of artificial intelligence problems. An algorithm of the inverse method application for such type formulas is justified. Upper and lower bounds of the number of steps in such an application are obtain. Upper bounds coincide with those of other deduction algorithms, but the exhaustion is greatly reduced while construction particular derivation.*

***Keywords**: artificial intelligence, pattern recognition, predicate calculus, inverse method of S.Yu.Maslov, complexity theory.*

***ACM Classification Keywords**: I.2.4 ARTIFICIAL INTELLIGENCE Knowledge Representation Formalisms and Methods – Predicate logic, I.5.1 PATTERN RECOGNITION Models – Deterministic, F.2.2 Nonnumerical Algorithms and Problems – Complexity of proof procedures.*

Introduction

One of the most thought out derivation methods for predicate calculus is the inverse method proposed by S.Yu. Maslov [Orevkov, 2003]. In [Kosovskaya, 2011] it is shown that many of artificial intelligence problems are reduced to the proof of a special type formulas. A simplification of the inverse method for the formulas of such a type is presented in this paper.

Bounds of the number of steps for solving an artificial intelligence problems (namely the problem of logic-objective recognition) using the Maslov inverse method are considered in this paper. The problem is the following [Kosovskaya, Timofeev, 1985]

Let Ω be a collection of finite sets $\omega=\{a_1,\dots,a_k\}$ which will be called objects. Any subset τ of ω will be called its part. A partition $\Omega = \cup_{j=1}^{K}\Omega_j$ of the set Ω on K (possibly intersected) classes is done. A set of predicates $p_1,\dots,p_n$ characterizes the properties of and relationships between the elements of ω.

Logical description $S(\omega)$ of the object ω is a set of all true constant formulas of the type $p_i(\bar{\tau})$ or $\neg p_i(\bar{\tau})$ calculated for all possible parts τ of the object ω.

Here and below the notation $\bar{x}$ is used for a list of elements of a finite set x, corresponding to some permutation of its elements. The fact that the list $\bar{x}$ elements are the elements of y will be written in the form $\bar{x} \subseteq y$. In order to write that list of values for variables $\bar{x}$ that satisfy the formula $A(\bar{x})$ are different the notation $\exists \bar{x}_{\neq} A(\bar{x})$ will be used. instead of the formula

$$\exists x_1 \ldots \exists x_n \left(\&_{i=1}^{n-1} \&_{j=i+1}^{n} (x_i \neq x_j) \, \& \, A(x_1, \ldots, x_n) \right).$$

It is shown in [Kosovskaya, 2011] that the most Artificial Intelligence problems may be reduced to the proof of the formula of the type

$$S(\omega) \Rightarrow \exists \overline{x}_{\neq} A_j(\overline{x}), \qquad (1)$$

where $A_j(\overline{x})$ is an elementary conjunction of atomic formulas with predicates $p_1,\ldots,p_n$. This problem is known to be an NP-hard problem [Kosovskaya, 2007]. The proof of the sequent (1) is equivalent to the proof of the formula

$$(\& \, S(\omega)) \rightarrow \exists \overline{x}_{\neq} A(\overline{x}),$$

where $(\&S(\omega))$ is a notification for a conjunction of all formulas from the set $S(\omega)$.

This formula can be reduced using the equivalent transformations to formula of the form

$$\forall a_1,\ldots,a_k \exists x_1,\ldots,x_{n\,\neq} \left(\mathop{\&}_{i=1}^{\alpha} D_i(a_1,\ldots,a_k,x_1,\ldots,x_n) \right),$$

where D_i is $\vee \neg S(\omega) \vee P_{k_i}(\overline{x})$ and the notation $\vee\neg S\ (\omega)$ means a disjunction of negations of formulas of $S\ (\omega)$, and the solution of which will be considered in the paper.

The Maslov Inverse Method and its Modification

We modify the Maslov inverse method described in [Orevkov, 2003] for the formulas of the form

$$\forall a_1,\ldots,a_k \exists x_1,\ldots,x_{n\,\neq} \left(\mathop{\&}_{i=1}^{\alpha} D_i(a_1,\ldots,a_k,x_1,\ldots,x_n) \right). \qquad (2)$$

The original inverse method is formulated for the formulas of the form

$$\forall z_1,\ldots,z_k \exists (x_1,\ldots,x_n)_{\neq} \forall y_1,\ldots,y_m \left(\mathop{\&}_{i=1}^{\alpha} D_i(z_1,\ldots,z_k,x_1,\ldots,x_n,y_1,\ldots,y_m) \right),$$

special case of which are the formulas considered in this paper. In our case there are no variables which are universally quantified at the inner and, respectively, the operations associated with them should be omitted.

Variables universally quantified at the outer will be considered as constants.

Below the definitions without numbering are the definitions from [Orevkov, 2003]. For the considered case definitions in the paper are numbered.

Definition. $x \prec y$ if there exists a formula $D_i(a_1,\ldots,a_k,u_1,\ldots,u_n,d_1,\ldots,d_m)$ from a list of formulas Γ of the form $D_i(a_1,\ldots,a_k,u_1,\ldots,u_n,d_1,\ldots,d_m)$ such that x coincides with one of the variables u_i, and y coincides with one of the variables d_j.

Definition: List Γ of the formulas of the form $D_i(a_1,\ldots,a_k,c_1,\ldots,c_n,d_1,\ldots,d_m)$ is called **admissible** under the following conditions.

1. For each formula D_i from Γ variables $a_1,\ldots a_k,\ d_1,\ldots,d_m$ are mutually distinct.

2. Whatever be the formulas $D_i(a_1,\ldots,a_k,u_1,\ldots,u_n,d_1,\ldots,d_m)$ and $D_r(a_1,\ldots,a_k,v_1,\ldots,v_n,c_1,\ldots,c_m)$ from Γ if a variable $d_l\ (1\leq l\leq m)$ coincides with a variable $c_p\ (1\leq p\leq m)$ then $l = p$.

3. It is impossible to find a string of variables $x_1,\ldots,x_l$ ($l \geq 1$) appearing in the formulas of Γ such that $x_1 \prec x_2 \prec \ldots \prec x_l \prec x_1$.

Since a_k are constants and d_m are absent in (2) then the conditions 1. – 3. hold for any D_i. Hence we can formulate the following definition.

Definition 1. Any list *Г* of formulas of the form $D_i(a_1,...,a_k,c_1,...,c_n)$ is **admissible** for the formulas of the form (2).

Definition: An admissible list *Г* of formulas of the form $D_i(a_1,...,a_k,c_1,...,c_n,d_1,...,d_m)$ is an **F-set** under the following conditions.

1. Formulas of the list *Г* are not repeated.
2. Whatever be the formulas $D_i(a_1,...,a_k,u_1,...,u_n,d_1,...,d_m)$ and $D_r(a_1,...,a_k,v_1,...,v_n,c_1,...,c_m)$ from *Г* if the sets $d_1,...,d_m$ and $c_1,...,c_m$ have a common variable then D_r and D_i coincide.

Definition 2. If the formulas in a list *Г* of formulas of the form $D_i(a_1,...,a_k,c_1,...,c_n)$ are not repeated then this list is an **F-set**.

Definition: Let $\alpha_1,...,\alpha_l$ and $\beta_1,...,\beta_l$ be the lists of variables and constants that can coincide with each other and with the constants from the list $a_1,...,a_k$. Let us consider the system of equalities

$$\begin{cases} \alpha_1 = \beta_1 \\ \cdots \\ \alpha_l = \beta_l \end{cases}.$$

Let $u_1,...,u_p$ be a list without repetitions of all variables which differ from constants $a_1,...,a_k$ that appear in equalities of the system (3). The system (3) is called a **system of equations in the variables** $u_1,...,u_p$. Any set of variables $(\gamma_1,...,\gamma_p)$ such that after simultaneous replacement of variables $u_1,...,u_p$ by their values in the solution $(\gamma_1,...,\gamma_p)$ the left and right parts of each equation of the system coincide is called a **solution of system of equations** (3). We say that the **solution σ_1 absorbs the solution σ_2** if σ_2 may be received from σ_1 by replacing some of the basic variables by other variables. **The solution** of the system of equations (3) is called **universal** if it absorbs all the solutions of this system.

For the considered case this definition takes the form

Definition 3. Let $\alpha_1,...,\alpha_l$ be a list of constants from the list $a_1,...,a_k$ and $\beta_1,...,\beta_l$ be a list of some variables and constants from the same list $a_1,...,a_k$. Consider the system of equalities

$$\begin{cases} \alpha_1 = \beta_1 \\ \cdots \\ \alpha_l = \beta_l \end{cases}. \qquad (3)$$

Let $u_1,...,u_p$ be a list without repetitions of all variables included in the equation system (3). System (3) is called a **system of equations with variables** $u_1,...,u_p$. Any set of constants $(\gamma_1,...,\gamma_p)$ from the list $a_1,...,a_k$ such that after simultaneous replacement of variables $u_1,...,u_p$ by their values in the solution $(\gamma_1,...,\gamma_p)$ the left and the right parts of each equation of the system coincide is called a **solution of system of equations** (3). **The system of equations** (3) **has no solution** if the list $(\gamma_1,...,\gamma_p)$ has repetitions.

In the case considered in the paper two variables in systems of equations cannot be compared (since different variables may have only different values according to the formulation of the problem). So the definitions of an absorbing and universal solutions are not used in the paper.

Procedure 1 (identifying of variables with constants). Let *Γ* be a list of formulas of the form $D_i(a_1,...,a_k,u_1,...,u_n)$, S be a system of equations with variables $u_1,...,u_p$ as unknowns, *σ* be a solution of this system. The procedure for identifying of variables with constants in the list of *Γ* according to the system *S* consists in the replacing of variables $u_1,...,u_p$ by their values from the solution of *σ* in all formulas from the list *Γ*.

Procedure of identifying formulas. Let *Γ* be a list of formulas in the form $D_i(a_1,...,a_k,c_1,...,c_n,d_1,...,d_m)$ and includes formulas $D_i(a_1,...,a_k,u_1,...,u_n,d_1,...,d_m)$ and $D_r(a_1,...,a_k,v_1,...,v_n,c_1,...,c_m)$. The procedure of identifying of these formulas in the list *Γ* consists in application to the list *Γ* of procedure of identifying of variables according to the system of equations in variables

$$\begin{cases} u_1 = v_1 \\ \ldots \\ u_n = v_n \\ d_1 = c_1 \\ \ldots \\ d_m = c_m \end{cases}.$$

If the system has a solution then the determined procedure succeeds.

In the considered case there are no variables $d_1,...,d_m$ and $c_1,...,c_m$ and the procedure will not be considered, as in the considered case different variables have different values. Hence the system has no solution.

The procedure of transformation of a list of formulas to an F-set. Let *Γ* be a list of formulas of the form $D_i(a_1,...,a_k,c_1,...,c_n,d_1,...,d_m)$. The procedure of transformation of a list of formulas *Γ* to an F-set consists in the sequential execution of the following operations.

- Check whether the list is admissible. If it is so then go to the next item. Otherwise, the procedure defined ends without a result.
- Reduce all the repetitions of formulas in the list.
- Check whether the list is an F-set. If it is so then the processed list is the result of the determined procedure. Otherwise, proceed to the next step.
- Find such formulas $D_i(a_1,...,a_k,u_1,...,u_n,d_1,...,d_m)$ and $D_r(a_1,...,a_k,v_1,...,v_n,c_1,...,c_m)$ in the list that sets $d_1,...,d_m$ and $c_1,...,c_m$ have a common variable. Apply the procedure of identifying of these formulas. If it succeeds then go to the next step. Otherwise, the defined procedure ends without a result.
- Check if there is a repetition of formulas in the resulting list of the previous step. If so, then go to step 1. Otherwise, the defined procedure ends without a result.

In our case, the first point should not be performed as any list is admissible. The fourth point also should not be performed. The fifth one in the absence of the fourth one repeats the second, so the procedure of transformation of the list of formulas in F-sets takes the form.

Procedure 2 (transformation of list of formulas to an F-sets). Let *Γ* be a list of formulas of the form $D_i(a_1,...,a_k,u_1,...,u_n)$. The transformation of the list *Γ* to an F-set consists in deleting repetitions of formulas in this list.

The procedure of gluing of formulas in F-sets. Let *Γ* be an F-set and *Γ* includes formulas *A* and *B*. The procedure of gluing of formulas *A* and *B* in the list *Γ* is sequential execution of the following operations.

- Apply the procedure of identification of formulas *A* and *B*. If it succeeds then go to the next step. Otherwise, the defined procedure ends without a result.
- Apply to the list of formulas obtained in the previous step the procedure for transformation of a list of formulas to an F-set.

In our case, this procedure will not be used because there is not used the procedure of identifying of formulas (the first step). The second step (without execution of the first step) is fulfilled automatically.

The procedure of constructing a closed F-sets: Let $D_i(a_1,...,a_k,t_1,...,t_n,d_1,...,d_m)$ and $D_r(a_1,...,a_k,v_1,...,v_n,c_1,...,c_m)$ be formulas of the form $D_i(a_1,...,a_k,c_1,...,c_n,d_1,...,d_m)$ where all the variables $a_1,...,a_k$, $t_1,...,t_p$, $d_1,...,d_m$, $v_1,...,v_p$ and $c_1,...,c_m$ are different. Denote the first formula by means of *A* and the second one by *B*. Assume that *A* contains an atomic formula $P(t_1, ..., t_S)$ as a disjunct and *B* contains the negation of the formula $P(v_1, ..., v_S)$, where P is an s -ary predicate. The procedure for constructing a closed F-set according to the pairs of formulas *A*, $P(t_1, ..., t_S)$ and *B*, $\neg P(v_1, ..., v_S)$ is the sequential execution the following operations.

1 Apply the procedure of identifying of variables to the list $\langle A, B\rangle$ according to the system of equations in variables

$$\begin{cases} t_1 = v_1 \\ \dots \\ t_s = v_s \end{cases}.$$

2 If the identifying procedure is successful then use the procedure of transformation of the resulting list of formulas in F-set.

Procedure 3 (construction of a closed F-set for formulas of the form (2)). Let $D_i(a_1,...,a_k,t_1,...,t_n)$ and $D_r(a_1,...,a_k,v_1,...,v_n)$ be the formulas of the form (2), where $t_1,...,t_p$ and $v_1,...,v_n$ are variables and $a_1,...,a_k$ are constants. Denote the first formula by means of *A*, and the second one by *B*. We will assume that *A* contains atomic formula $P(t_1, ..., t_S)$ as a disjunct, and B contains the negation of the formula $P(a_1, ..., a_S)$, where *P* is an s-ary predicate. The procedure for constructing a closed F-sets of pairs of formulas *A*, $P(t_1, ..., t_S)$ and *B*, $\neg P(a_1,..., a_S)$ is the sequential execution of the following operations.

1. Apply the procedure of identifying of variables to the list $\langle A, B\rangle$ according to the system of equations in variables

$$\begin{cases} t_1 = a_1 \\ \dots \\ t_s = a_s \end{cases}.$$

2. If the identifying procedure is successful then use the procedure of transformation of the resulting list of formulas in F-set.

Definition 4: F-set is called a **closed** one if it may be obtained by applying the procedure of constructing a closed F-set with some pairs of formulas *A, P* $(t_1, ..., t_s)$ and *B*, $\neg P(a_1,..., a_s)$.

Procedure of application of Rule B to an F-set. Consider a system of δ F-sets

$$\begin{cases} \Gamma_1, \quad D_1(a_1,...,a_k,c_1^1,...,c_n^1,d_1^1,...,d_m^1) \\ \quad ... \\ \Gamma_\delta, \quad D_\delta(a_1,...,a_k,c_1^\delta,...,c_n^\delta,d_1^\delta,...,d_m^\delta) \end{cases}, \tag{4}$$

where $\Gamma_1,\ldots,\Gamma_\delta$ are the lists of formulas of form $D_i(a_1,...,a_k,c_1,...,c_n,d_1,...,d_m)$. If any two of these sets contain a common basic variable then rename one of them to a new variable. In such a way we will achieve that F-sets (4) will not contain common basic variables. Let us apply procedure of identifying of the variables of the lists $\Gamma_1,\ldots,\Gamma_\delta$ according to the system of equations in variables

$$\begin{cases} c_1^1 = c_1^2 = ... = c_1^\delta \\ \quad ... \\ c_n^1 = c_n^2 = ... = c_n^\delta \\ d_1^1 = d_1^2 = ... = d_1^\delta \\ \quad ... \\ d_m^1 = d_m^2 = ... = d_m^\delta \end{cases} \tag{5}$$

and then the procedure of transformation of list of formulas in F-sets to the resulting list. Constructed in such a way F-set Σ is the result of application of the rule B to F-sets (4), if the values of variables $d_1^1, d_2^1, ..., d_m^1$ in the universal solution σ of (5) satisfy the following conditions.

1. They are distinct.
2. They are not included in the formula of Σ.
3. They differ from the values $c_1^1, c_1^2, ..., c_1^\delta$,..., $c_n^1, c_n^2, ..., c_n^\delta$ in the solution variables σ.

As in our case, there are no variables $d_1^1, d_1^2, ..., d_1^\delta$, ..., $d_n^1, d_n^2, ..., d_n^\delta$ this rule has the following form:

Procedure 4 of application of Rule B to F-sets. Consider a system of δ F-sets

$$\begin{cases} \Gamma_1, \quad D_1(a_1,...,a_k,c_1^1,...,c_n^1) \\ \quad ... \\ \Gamma_\delta, \quad D_\delta(a_1,...,a_k,c_1^\delta,...,c_n^\delta) \end{cases}, \tag{6}$$

where $\Gamma_1,\ldots,\Gamma_\delta$ are lists of the formulas of form $D_i(a_1,...,a_k,u_1,...,u_n)$. If any two of these sets contain a common variable then rename it to a new variable. In such a way it will be achieved that F-sets (6) do not contain common variables. Let us apply procedure of identifying of the variables to the list of $\Gamma_1,\ldots,\Gamma_\delta$ according to the system of equations in the variables

$$\begin{cases} c_1^1 = c_1^2 = ... = c_1^\delta \\ \quad ... \\ c_n^1 = c_n^2 = ... = c_n^\delta \end{cases} \tag{7}$$

and then procedure of transformation of list of formulas in F-sets must be applied to the resulting list. Constructed in this way F-set *Σ* is the result of application of rule B to F-sets (6).

Since at the very beginning of the first application of this rule we have renamed all the variables then all variables in the system of equations (7) are different. So contrary to the condition that different variables have different values in the solution of the system does not arise. While solving equations (7) the variables are renamed again to the original names and *δ* F-sets are combined into a single F-set so the application of the rule B to the formulas considered in this paper can be reduced to a simple unification of δ monomial F-sets in a δ-member F-set. For simplicity, we will use this procedure in such a way: we should rewrite one δ-member F-set instead of δ monomial F-sets.

Theorem [Orevkov, 2003]**.** The formula F is provable in a predicate calculus if and only if an empty F-set □ is derivable in the calculus of favorable sets.

This calculus is given by following **rules A, B and the rule of permutation**

Rule A: Closed F-set is favorable.

Rule B: F-set is favorable if the procedure of rule B applying is successful.

Rule of permutation: A permutation of formulas in a favorable F-set is a favorable F-set.

Now, on the basis of the presented procedures and rules we can formulate the algorithm for searching an inference of a formula of the form (2)

$$\forall a_1,...,a_k \exists (x_1,...,x_n)_{\neq} \left(\underset{i=1}{\overset{\alpha}{\&}} D_i(a_1,...,a_k,x_1,...,x_n) \right)$$

using the tactic of the inverse method and compare it with the algorithm of proof search using tactics of the resolution method.

Algorithm of Formula Derivation Based on Maslov Inverse Method

Definition 5. F-set is called **empty** if all formulas in it have no variables and are tautological.

Definition 6. F-set is called a **deadlock one** if it includes at least one formula that has no variables and is false or it is neither a tautology nor a contradiction.

Algorithm *Alg* of the formula $\forall a_1,...,a_k \exists (x_1,...,x_n)_{\neq} \left(\underset{i=1}{\overset{\alpha}{\&}} D_i(a_1,...,a_k,x_1,...,x_n) \right)$ **proof .**

1. F-sets $\{D_1(a_1,...,a_k, x_1,...,x_n), ..., D_\alpha(a_1,...,a_k, x_1,...,x_n)\}$ are written down according to the original formula.
2. Apply the procedure for constructing a closed F-set as follows:

   2.1. We are looking for such elementary disjunctions D_j and D_r which contain *s*-ary predicate symbol *P* and its negation (with may be different sets of variables or constants as arguments). Let it be an atomic formula $P(t_1, ..., t_S)$ and the negation of atomic formula $\neg P(v_1, ..., v_S)$. One of the lists $t_1, ..., t_S$ or $v_1, ..., v_S$ must be a list only of constants. If the other list contains a constant then the two lists at the same position should have the same constant. For example, if one of the lists has the form (a_2, a_4, a_3, a_1) then the other one must contain the constant a_2 or a variable on the first place, the constant a_1 or a variable on the second place, and so on.

   2.2. Solve the system of equations which identifies the lists of variables and constants $t_1,...,t_s$ and $v_1,...,v_s$. If this system has a solution then go to step 2.3. Otherwise go to step 2.1. as follows: look for another version appropriate for formula or circuit considered suitable for the new closed formula.

2.3. Delete repetition of formulas (if they exist) in the resulting F-set.

2.4. Verify whether an empty set is derived. If it is so, the algorithm run stops. Otherwise, if there exist a formula in the F-set to which a rule for constructing a closed F-set can be applied then go to step 3. If a deadlock set is received then go to step 3.

3. Undo one action, and execute step 2. as follows: look for another version appropriate for formula or circuit considered suitable for the new closed formula. If the application of step 2. does not give the new assign values to variables, step 3 is applied once more. Apply rule 3. until the empty set is derived or an opportunity to cancel the action runs out.

4. If the combination for the closure of F-sets is over, and the result is not obtained, then the formula is not derivable

Estimates of the Number of Steps of Algorithm Run

Let $\forall a_1,\ldots,a_k \exists (x_1,\ldots,x_n)_{\neq} \left(\underset{i=1}{\overset{\delta}{\&}} D_i(a_1,\ldots,a_k,x_1,\ldots,x_n) \right)$ be a considered formula, where k is a number of constants; n be the number of variables ($k \geq n$ since values of different variables must differ); D_i has the form $\vee \neg S(a_1,\ldots,a_k) \vee P_{k_i}(\bar{x})$ (where $\vee\neg S$ *(ω)* means a disjunction of negations of formulas of *S (ω)*) and the solution of which will be considered in the paper.

The execution of every of the following operations is taken for one **step**:

- assignment of a variable value (the solution of an equation of the form *x* = *a*);
- verifying the graphic coincidence of atomic formulas;
- substitution of a variable value into a formula.

Introduce the following notations:

l is the maximal number of arguments in the atomic formula;

s is the number of atomic formulas in *S (ω);*

γ_l *is* the number of *i*-ary predicates in *S (ω) ;*

γ_i^0 *is* the number of *i*-ary anomic formulas in *S (ω)* without negation;

$\gamma_i^{\neg}$ *is* the number of *i*-ary anomic formulas in *S (ω)* with negation.

Theorem 1. (The lower bound of the algorithm run.) The number of steps of the proof of the formula $\forall a_1,\ldots,a_k \exists (x_1,\ldots,x_n)_{\neq} \left(\underset{i=1}{\overset{\delta}{\&}} D_i(a_1,\ldots,a_k,x_1,\ldots,x_n) \right)$ with the use of an algorithm based on the tactics of the inverse method is not less then *sl*.

Proof. The lower bound is achieved when the number of variables is equal to the maximal number of arguments in the atomic formula and the answer is obtained by solving a system of *l* equations. Every elementary disjunction contains *s* constant formulas. Hence we have not less then *sl* steps. ■

Let $\tilde{s}$ be the maximal number of occurrences of the same predicate (only without the negations, or only with the negations) in the class description.

Theorem 2. (The upper bound of the algorithm run.) The number of steps of the proof of the formula $\forall a_1,...,a_k \exists (x_1,...,x_n)_{\neq} \left(\mathop{\&}\limits_{i=1}^{\delta} D_i(a_1,...,a_k,x_1,...,x_n) \right)$ with the use of an algorithm based on the tactics of the inverse method is $O(\delta \tilde{s}^{\delta} l)$.

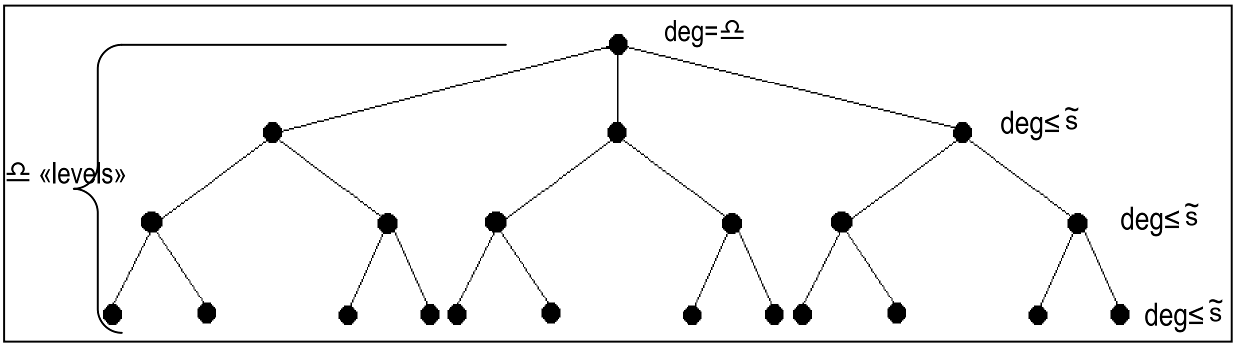
Proof. The first F-set under consideration is a list of the form $D_1(a_1,...,a_k, x_1,...,x_n), ..., D_\alpha(a_1,...,a_k, x_1,...,x_n)$, where every $D_j(a_1,...,a_k,x_1,...,x_n)$ is an elementary disjunction of the type $\vee \neg S(a_1,...,a_k) \vee P_{k_i}(\bar{x})$ with the same constant formulas.

Let the atomic formulas in $S(a_1,...,a_k)$ are ordered by groups with the same predicate according decreasing of number of arguments. In every group predicates without negation precedes ones with negation.

If *p(l)* denotes some atomic formula with constant arguments, *p(n)* denotes some atomic formula with variable arguments then the list has the following structure:

$$\left.\begin{array}{l} \mathop{\vee}\limits_{i_1=1}^{\gamma_l^0} p_{i_1}(\bar{a}_{i_1}) \vee \mathop{\vee}\limits_{i_2=1}^{\gamma_l^{\neg}} \neg p_{i_2}(\bar{a}_{i_2}) \vee \mathop{\vee}\limits_{i_3=1}^{\gamma_{l-1}^0} p_{i_3}(\bar{a}_{i_3}) \vee \mathop{\vee}\limits_{i_4=1}^{\gamma_{l-1}^{\neg}} \neg p_{i_4}(\bar{a}_{i_4}) ... \vee \mathop{\vee}\limits_{i_{2l-1}=1}^{\gamma_1^0} p_{i_{2l-1}}(\bar{a}_{i_{2l-1}}) \vee \mathop{\vee}\limits_{i_{2l}=1}^{\gamma_1^{\neg}} \neg p_{i_{2l}}(\bar{a}_{i_{2l}}) \vee p_1(\bar{x}_1) \\ \qquad \qquad \qquad \qquad ... \\ \mathop{\vee}\limits_{i_1=1}^{\gamma_l^0} p_{i_1}(\bar{a}_{i_1}) \vee \mathop{\vee}\limits_{i_2=1}^{\gamma_l^{\neg}} \neg p_{i_2}(\bar{a}_{i_2}) \vee \mathop{\vee}\limits_{i_3=1}^{\gamma_{l-1}^0} p_{i_3}(\bar{a}_{i_3}) \vee \mathop{\vee}\limits_{i_4=1}^{\gamma_{l-1}^{\neg}} \neg p_{i_4}(\bar{a}_{i_4}) ... \vee \mathop{\vee}\limits_{i_{2l-1}=1}^{\gamma_1^0} p_{i_{2l-1}}(\bar{a}_{i_{2l-1}}) \vee \mathop{\vee}\limits_{i_{2l}=1}^{\gamma_1^{\neg}} \neg p_{i_{2l}}(\bar{a}_{i_{2l}}) \vee p_\delta(\bar{x}_\delta) \end{array}\right\} \delta \quad \text{items}$$

The upper bound is reached, if the answer whether a formula is derivable is received at the last step of the algorithm. That is if you are to avoid all the branches of a tree the following form:



The number of levels in the tree equals to δ as it is necessary to assign values to variables in δ formulas. δ edges leaves from the upper node. Not more than $\tilde{s}$ edges leave from each node of the next levels. The number of nodes in the tree is not more than $\delta(\tilde{s}^{\delta}-1)$. Every node corresponds to a system of not more than *l* equations every of which may be verified not more than in *l* steps. That is the upper bound is $O(\delta \tilde{s}^{\delta} l)$, where δ is the number of disjunctive terms in the original formula, $\tilde{s}+1$ is the total number of atomic formulas in every disjunct, *l* is the largest number of arguments in the atomic formulas. ■

Conclusion

Thus, the asymptotic estimation of the number of steps of the described algorithm *Alg* using the tactics of the inverse method are as follows

$$O(sl) \leq T(Alg) \leq O(\delta \tilde{s}^{\delta} l),$$

where *s+1* is the total number of atomic formulas in disjuncts, δ is the number of disjunctive terms in the original formula, *l* - the largest number of arguments in the atomic formulas and $\tilde{s}$ is the maximal number of occurrences of the same predicate (only without the negations, or only with the negations) in the class description.

In [Kosovskaya, 2007] the following upper bound for solving the pattern recognition algorithm that uses the tactics of the method of resolution

$$O(D\,\tilde{s}^{\tilde{a}})$$

where D is the number of disjuncts in the descriptions of the classes used in problem solving,

$\tilde{s}$ is the maximum number of occurrences of the same predicate (only without the negations, or only with the negations) in the class description, $\tilde{a}$ is the maximum number of occurrences of atomic formulas of the elementary conjunctions that make up the class definitions was obtained.

As we can see, these estimates coincide to within a constant factor.

Bibliography

[Kosovskaya, Timofeev, 1985] T.M. Kosovskaya, A.B. Timofeev. About one new approach to the formation of logical decision rules in pattern recognition problems // Vestnik LGU, 1985, №8, pp. 22-29. (In Russian)

[Kosovskaya, 2007] Kosovskaya T.M. Proofs of the number of steps bounds for solving of some pattern recognition problems with logical description // Vestn. S- Petersburg University, Ser. 1, 2007. Ed. (2), pp. 82-90. (In Russian)

[Kosovskaya, 2011] Kosovskaya T. Discrete Artificial Intelligence Problems and Number of Steps of their Solution // International Journal "Information Theories and Applications", Vol. 18, Number 1, 2011. P. 93 – 99.

[Orevkov, 2003] Orevkov V.P. The inverse method of logical derivation // In Adamenko A. Kuchukov A. Programming Logic and Visual Prolog. St. Petersburg, "`BHV-Petersburg"'. (2003), pp.. 952-965. (in Russian)

Authors' Information

Tatiana Kosovskaya *– Dr., Senior researcher, St.Petersburg Institute of Informatics and Automation of Russian Academy of Science, 14 line, 39, St.Petersburg, 199178, Russia; Professor of St.Petersburg State Marine Technical University, Lotsmanskaya ul., 3, St.Petersburg, 190008, Russia; Professor of St.Petersburg State University, University av., 28, Stary Petergof, St.Petersburg, 198504, Russia, e-mail: kosov@NK1022.spb.edu*

Major Fields of Scientific Research: Logical approach to artificial intelligence problems, theory of complexity of algorithms.

Nina Petukhova *– PHD student, St.Petersburg State Marine Technical University, Lotsmanskaya ul., 3, St.Petersburg, 190008, Russia, e-mail: ninka_mat@mail.ru*

Major Fields of Scientific Research: Logical approach to Artificial Intelligence problems.

The paper is published with financial support of the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

POLYNOMIAL-TIME EFFECTIVENESS OF PASCAL, TURBO PROLOG, VISUAL PROLOG AND REFAL-5 PROGRAMS

Nikolay Kosovskiy, Tatiana Kosovskaya

***Abstract**: An analysis of distinctions between a mathematical notion of an algorithm and a program is presented in the paper. The notions of the number of steps and the run used memory size for a Pascal, Turbo Prolog, Visual Prolog or Refal-5 program run are introduced. For every of these programming languages a theorem setting conditions upon a function implementation for polynomial time effectiveness is presented.*

*For a Turbo or Visual Prolog program It is proved that a polynomial number of steps is sufficient for its belonging to the class **FP**. But for a Pascal or Refal-5 program it is necessary that it additionally has a polynomially bounded run memory size.*

***Keywords**: complexity theory, class **FP**, programming languages Pascal, Turbo Prolog, Visual Prolog and Refal-5.*

***ACM Classification Keywords**: F.2.m ANALYSIS OF ALGORITHMS AND PROBLEM COMPLEXITY Miscellaneous.*

Introduction

What is an effectively in time computable function? In the frameworks of computational complexity it is accepted that this is a function which belongs **FP**. The definition of this class is done in the terms of a Turing machine. The class **FP** is defined as the class of all algorithms which may be calculated by a Turing machine in a polynomial under the input word length number of steps [Du 2000]. Theoretically the Turing machine is a very good instrument because all the steps fulfilled by it while a function computation are identically simple and may be regarded as using the same amount of time. But who writes a Turing machine program? We prefer a high level programming language.

Can we estimate a run time of a program according to the number of the fulfilled operators? For example, is it true that the fulfilling of the next tree operators needs three units of time?

x:=z^2; y:=5; z:=log(sin($x*y$));

In dependance of a translator implementation the first operator either may be represented in the form *if* <exponent> = 2 *then x:=z*z* or the value of *x* is calculated be the formula e^{2*ln(z)}. The values of logarithm and exponent are calculated with some (enough good) precision. The first alternative is used in the most of translators for calculation of a square, cube, ... And what is for 25th power?

The second operator, of course, uses one unit of time.

The calculation of the third operator as a rule uses an expansion in series of functions log and sin.

And how many steps of calculation are here?

It is well known that the solving of a system of linear equations with integer coefficients by means of Gauss method may be implemented in a polynomial under the dimensionality of the system number of operations (addition and multiplication). But even if we solve a system with 5 variables with great enough absolute values of coefficients then the overflow occurs. The “fighting” with such an overflow demands the necessity either to introduce a structure simulating a number of great capacity or to use an approximate method. But can we guarantee that even after that the number of the used operations is adequate to the time of a program run?

Below there are presented theorems setting conditions under which a program in Pascal, Turbo Prolog, Visual Prolog or Refal-5 [Babaev 1992] performing a polynomial number of steps calculate a function belonging to the class **FP**. Such a program may be called a polynomially effective one.

Another approach to the simulation of the class **FP** is presented in [Beltiukov 2006]. This approach is based on the selection of such a subset of Pascal which provides to receive such a class of functions which equals to **FP**. It uses a variant of bounded recursion introduced by A. Grzegorczyk.

A program as an algorithm

Some researchers assume that the notions of a program and an algorithm coincide. Somebody suppose that a program is an implementation of an algorithm. Let's try to clear up the question. Here under the notion “algorithm” we will understand a mathematical notion of an algorithm such as a Turing machine, Markov normal algorithm, Markov-Post algorithm [Kosovskaya 2010]. Here under the notion “program” we will understand a program written in such a language as one from the Pascal languages family, or such programming languages destined for the solving of traditional Artificial Intelligence problems as Turbo Prolog, Visual Prolog and Refal-5 [Babaev 1992].

A mathematical notion of an algorithm implies a potentially infinite set of both the input data and the results of an algorithm run.

That is why only file may be used as an input data for a language from the Pascal languages family regarded as a mathematical notion of an algorithm. But the notion of the value type *integer* may be extended in the following way. We allow to use integers with unbounded number of digits. Such an unbounded number may be implemented with the use of dynamic linear arrays with the elements of the initial value type *integer*. A function of such a new type integers will be called a pascal-like function. The set of all such functions for every language from the Pascal languages family may be regarded as a mathematical notion of an algorithm.

Analogous notions may be introduced for the other programming languages from the Pascal languages family.

For the dialects of Prolog it may be, for example, lists of elements of the value type *integer* or files or constant terms without variables in some (not interpreted) signature.

Here a mathematical notion of transformation object for Turbo or Visual Prolog language is regarded as any list of elements of the value type *integer* (including an empty list). As a rule such an element is a remainder modulo 2^{16}, more precisely any integer between -2^{15} and $2^{15} - 1$ inclusively. In such a case Turbo or Visual Prolog may be regarded as a mathematical notion of an algorithm.

For the programming language Refal-5 analogous notion of transformation object may be a sequence of macro-digits which are separated in the program text by blanks. A macro-digit is a digit 0 or a sequence of decimal digits not beginning with 0 and defining a decimal notation of a positive integer less then 2^{32}.

Number of steps and run memory size for some algorithm notions

It is accepted in the theory of algorithm complexity that every run complexity characteristic is regarded as a function of the input data length. This function equals maximum of the complexity characteristic under consideration upon all input data with the same length. The value of such a characteristic will be estimated up to a multiplicative constant. The notion of the input data for every programming language under consideration was done in the previous section.

The **number of steps** and the **run memory size** are the most naturally defined **for a Turing** machine. They are respectively the number of fulfilled commands and the number of cells visited by a Turing machine head.

The **number of steps** and the **run memory size for the normal Markov algorithms** may be defined respectively as the number of fulfilled substitutions and the maximal word length in the sequence of transforming words beginning with the input word and ending with the result of algorithm run.

The **number of steps for a pascal-like function** is defined as the number of fulfilled statements and computed expressions and sub-expressions (boolean and arithmetic) during the run of a program receiving a result according to the rules of Pascal semantics.

The **run memory size for a pascal-like function** is defined as maximum upon all computational steps (beginning with the input data and ending with the result) of the sums of the record lengths of all calculated variables values (including the values of all elements of all arrays) . The length of the stack of variable values for fulfilling embedded procedures and pascal-like functions (including recursive ones) is taken in account. While running a call of such a procedure or function a separator followed by actual parameters and all local variable values of the call is put into this stack. Besides that the length of the stack of definition descriptions of the called procedures and functions sufficient for all program run is taken in account. The notation length of a variable or an array element which has no value up to that moment equals 1.

As you can see the full definition of the number of steps and the run memory size for a pascal-like function contains full definition of Pascal operational semantics for its used version.

The **number of steps for a Turbo or Visual Prolog query** is defined as the number of fulfilled calls of a predicate appeared during the run of this query (including all calls of built in predicates). It is clear that this definition is much shorter than the definition of number of steps of pascal-like function.

The **number of steps for a Refal-5 function** is the number of fulfilled Refal-5 rules.

The **run memory size for a Refal-5 function** is the sum of the length of expression and the length of a built-in Refal-5 stack, maximal upon all steps of calculation. I.e. the maximum of the lengths of all expression records (beginning from the initial and ending the final) appearing according to the semantics of refal-5 functional expressions computation.

The number of steps and the run memory size for the regarded algorithm notions will be used in the theorems of the next section.

Polynomial-time computations

FP is the usual denotation of the class of all functions computable by a Turing machine with the number of steps not more than a polynomial under the length of the input word [Du 2000].

The words "Turing machine" may be replaced by the words "normal Markov algorithm" (as it was proved by G.S. Tseitin in the Leningrad seminar on mathematical logic approximately 40 years ago). He proved the following theorem.

Theorem 1. The class of all functions computable by a normal Markov algorithm with the number of steps not more than a polynomial under the length of the input word equals to the class **FP**.

The proof of this theorem is based on the simulation of a normal Markov algorithm by a Turing machine with a suitable number of steps.

This analogous theorem is not valid for pascal-like functions. For example, the function 2^n may be calculated by a pascal-like function in not more than *log(n)* (which is approximately equal to the length of *n*) number of steps. But this function does not belong to **FP** as the length of its result is n and hence the number of a Turing machine steps can't be less than $n \approx 2^{||n||}$, where $||n||$ denotes the value length of n. This function is exponential of input data length.

Let the definition of a pascal-like function does not contain pointers, sets, records and files as well as procedures and functions as parameters.

A pascal-like function is called **double polynomial** if both the number of steps and the run memory size are less than a polynomial under the record length of input data.

Theorem 2. [Kosovskaya 2010] The class of all double polynomial pascal-like functions equals to the class **FP**.

It means that for an effective pascal-like function it is not sufficient only a polynomial number of steps.

The above formulated theorem has the longest proof among the theorems formulated in this paper.

Let a Turbo or Visual Prolog program has no calls of the built-in predicate *concat* or of such a predicate which processes a file or a term with an uninterpreted function. Usually such a term is used as a record of a tree.

Theorem 3. The class of all functions computable by queries situated in the goal section of a Turbo or Visual Prolog program in a polynomial under summary length of input data number of steps equals to the class **FP**.

The proof of this theorem is sufficiently shorter than any proof of other theorems from this paper except theorem 1.

For functions computable by a Refal-5 program the situation is the same as for pascal-like functions.

Let any condition for application of a rule in the definition of a Refal-5 function does not contain a recursive call of the defining function. And let the definition of a Refal-5 function does not contain commands processing files.

A Refal-5 function is called **double polynomial** if both the number of steps and the run memory size are less than a polynomial under the length of its input expression.

Theorem 4. [Kosovskiy 2011] The class of all double polynomial Refal-5 functions equals to the class **FP**.

The proof of this theorem is much shorter than that of theorem 2 but essentially longer than the one of theorem 3.

Conclusion

It is well known the question whether an NP-complete problem may be computed by a Turing machine in a polynomial number of steps [Garey 1979]. The presented theorems allow to replace in such a question the notion of Turing machine by more practically used notions of programming languages.

For every regarded programming language it is possible to introduce such a complexity measure that a polynomial under the input data length of which provides the belonging of a function computable by a program to the class **FP**.

The shortest conditions were received for Turbo or Visual Prolog programs. Polynomial in time computation by a Turbo or Visual Prolog program equals to the same type computation by a Turing machine.

For more widespread programming languages such as Pascal and Refal-5 checking out the polynomial in time effectiveness needs to take in account not only the number of steps but also the run memory size.

Bibliography

[Babaev 1992] I.O.Babaev, M.A.Gerasimov, N.K.Kosovskiy, I.P.Soloviev. Intellectual programming. Turbo Prolog and Refal-5 for personal computers. St.Petersburg State University Press. 1992. (In Russian)

[Beltiukov 2006] A.P. Beltiukov , A Polynomial Programming Language. In: "Mathematical Problems of Computer Science", Transactions of the Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia, , vol.27, 2006.

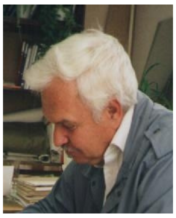
[Du 2000] D.Z.Du, K.I.Ko Theory of Computational Complexity. A Willey-Interscience Publication. John Wilet & Sons, Inc. 2000.

[Garey 1979] M.R.Garey, D.S.Johnson. Computers and Intractability. A Guide to the theory of NP-Completeness. Bell Laboratories. Murray Hill, New Jersey. W.H.Freeman and Company, San Francisco, 1979.

[Kosovskaya 2010] T.M.Kosovskaya, N.K.Kosovskiy. Bases of Polynomial-Time Proof for the Simplest Mathematical Algorithms. In: Computer Instruments in Education. No 2, 2010. (In Russian)

[Kosovskiy 2011] N.K.Kosovskiy. Church Thesis for Polynomial in Time Recursive Algorithms in Words and their Lengths. In: Computer Instruments in Education. No 1, 2011. (In Russian)

Authors' Information



Nikolay Kosovskiy *– Dr., Professor, Head of Computer Science Chair of St.Petersburg State University, University av., 28, Stary Petergof, St.Petersburg, 198504, Russia, e-mail: kosov@NK1022.spb.edu*

Major Fields of Scientific Research: Mathematical Logic, Theory of Complexity of Algorithms.

Tatiana Kosovskaya *– Dr., Senior researcher, St.Petersburg Institute of Informatics and Automation of Russian Academy of Science, 14 line, 39, St.Petersburg, 199178, Russia; Professor of St.Petersburg State Marine Technical University, Lotsmanskaya ul., 3, St.Petersburg, 190008, Russia; Professor of St.Petersburg State University, University av., 28, Stary Petergof, St.Petersburg, 198504, Russia, e-mail: kosov@NK1022.spb.edu*

Major Fields of Scientific Research: Logical Approach to Artificial Intelligence Problems, Theory of Complexity of Algorithms.

The paper is published with financial support of the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

TABLE OF CONTENT