ITHEA

International Journal INFORMATION THEORIES & APPLICATIONS

2014 Volume 21 Number 2

International Journal
INFORMATION THEORIES & APPLICATIONS
Volume 21 / 2014, Number 2

International Journal "INFORMATION THEORIES & APPLICATIONS" (IJ ITA) is official publisher of the scientific papers of the members of the ITHEA International Scientific Society

IJ ITA welcomes scientific papers connected with any information theory or its application.
IJ ITA rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITA are given on www.ithea.org.

Responsibility for papers published in IJ ITA belongs to authors.
General Sponsor of IJ ITA is the **Consortium FOI Bulgaria** (www.foibg.com).

International Journal "INFORMATION THEORIES & APPLICATIONS" Vol. 21, Number 2, 2014

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 Politécnica de Madrid, Spain,
Hasselt University, Belgium,
St. Petersburg Institute of Informatics, RAS, Russia,
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia

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

ISSN 1310-0513 (printed) **ISSN 1313-0463 (online)** **ISSN 1313-0498 (CD/DVD)**

PROGRAMMING OF AGENT-BASED SYSTEMS

Dmitry Cheremisinov, Liudmila Cheremisinova

Abstract: *The purpose of the paper is to explore the possibility of applying the language PRALU, proposed for description of parallel logical control algorithms and rooted in the Petri net formalism for design and modeling real-time multi-agent systems. It is demonstrated with a known example of English auction on how to specify an agent interaction protocol using considered means. A methodology of programming agents in multi-agent system is proposed; it is based on the description of its protocol on the language PRALU of parallel algorithms of logic control. The methodology consists in splitting agent programs into two parts: the block of synchronization and the functional block.*

Keywords: *multi-agent system, interaction protocol, BDI agent, parallel control algorithm.*

ACM Classification Keywords: *I.2.11 [Computer Applications]; Distributed Artificial Intelligence, Multiagent systems; D.3.3 [Programming Languages]: Language Constructs and Features – Control structures, Concurrent programming structures*

Introduction

In recent years one of the most rapidly growing areas of interest for distributed computing is that based on the concept of agents and multi-agent systems (MAS). The first MAS applications have appeared in the mid-1980s. Now they are becoming one of the most important topics in distributed and autonomous decentralized systems. There are increasing attempts to use agent technologies in variety of domains, ranging from manufacturing to process control, air traffic control and information management. Programming technology based on the use of interacting agents is considered the most promising tool of modern programming.

MAS is a computational system which consists of a number of agents that operate together in order to perform a set of tasks or to achieve a set of goals [Burmeister, 1992; Jennings, 2001; Lesser, 1999; Subrahmanian, 2000]. There exists variety of definitions of the notion of agent. The most of researchers adhere to the definition of M. Wooldridge and N.R. Jennings: agent is a hardware or software-based computer system that enjoys the following properties:

– *Autonomy*: agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal states;

– *Social ability*: agents interact with other agents (and possibly humans) via some kind of *agent-communication language*;

– *Reactivity*: agents perceive their environment, and respond in a timely fashion to changes that occur in it;

– *Pro-activeness*: agents do not simply act in response to their environment, they are able to exhibit goal-directed behavior by taking the initiative.

MAS are usually specified as a concurrent system that consists of autonomous, reactive and internally-motivated agents acting in a decentralized environment. One key reason of the growth of interest in MAS is that the idea of an agent as an autonomous system, capable of interacting with other agents, is a naturally appealing one for software designers. Reactive agents do not have representations of their external environment and act using a stimulus response type of behavior, they respond to the present state of the environment. Robots are examples of artificial agents. ARCHON [Burmeister, 1992] is the most well-known commercial system constructed on the basis of the conception of agents and designed for controlling the process of manufacturing articles.

The history of the development of the agent theory began with the problem of modeling the properties of living systems and dates back to the work of W. Pitts and W. MacCulloch on formal neurons, John von Neumann on self-reproducing automata, A. N. Kolmogorov on the theory of complexity, H. von Forster and Ilya Prigogine on the theory of self-organization, William Ashby on models of homeostasis, W.G. Walter on reactive robots, and John Holland on genetic algorithms.

Designing and building agent systems is difficult. They have all the problems associated with building traditional distributed, concurrent systems and have the additional difficulties that arise from having flexible and sophisticated interactions between autonomous problem-solving components. The big question then becomes how effective MASs can be designed and implemented.

The core concept of multi-agent systems is interaction; it is the foundation for cooperative or competitive behavior among several autonomous agents. Agent interactions are established through exchanging information in the form of messages that specify the desired performatives of interacting agents. Formalization of the concept of interaction as a method of transmitting messages from one dispatcher to several recipients through a transmission environment derives from Claude Shannon’s work on communication theory [Shannon, 1948]. The subsequent formalization of this concept employed the theory of speech acts [Searle, 1986]. In this theory communication between agents is considered as a form of behavior, since particular types of sentences of natural language are of the nature of actions (called speech acts) and presume a “rational effect.” Speech acts form the basis of the communication languages used by agents (e.g., the KQML and FIPA-ACL languages [ARPA, 1993; Finin, 1997; FIPA, 2002]) these languages also define the sets of permissible speech acts and semantics associated with these sets.

A set of interrelated messages forms conversations in which agents play different roles as a functions of their individual or common goals. Conversations in multi-agent systems are based on interaction protocols that define all possible flows of conversations. Agent system can operate if agents have a common understanding of the possible types of messages, then they must know which messages they can expect in a particular situation and what they may do when they got some message. So messages exchanged between agents in some multi-agent system need to follow some standard patterns which are described in agent interaction protocol. Protocols play a central role in agent communication; they specify the sequences in which messages should be arranged in agent’s interactions.

At this time, there are two major technical obstacles to the widespread adoption of multiagent technology: 1) the lack of systematic methodologies enabling designers to clearly specify and structure their applications as MASs and 2) the lack of widely available MAS toolkits. Flexible sets of tools are needed that enable designers to specify an agent’s problem-solving behavior and to specify agents interaction, and then visualize and debug the behavior of the agents and the entire MAS. That is because designing agent systems is difficult enough: they have all the problems associated with building traditional distributed, concurrent systems and have the additional difficulties that arise from having sophisticated interactions between autonomous problem-solving components.

The other difficulty of existing methodologies of MAS designing consists in follows. In the majority of MAS models autonomous behavior of agents is described in terms of belief, desires and intentions (BDI) [Burmeister, 1992], and their communications are specified in terms of protocols which have no a direct relation with the first formalism. The behavior of agents is described in terms of formalisms of a high level abstraction (such as temporal logic), but the communication is specified in the concepts close to realization. The independent behavior of agents in the majority of models of multi-agent systems is described by means of formalisms of high level abstractness, but the communication is specified by the concepts close to realization. The difference of levels of the description does not allow model communications between agents at the level in which their independent autonomous behavior is described. This problem arises because of absence of agent models that unify all aspects of local behavior and the communications (BDI-behavior and communication), and will be suitable for automated implementation. The main reason for the absence of such a unifying model is that there is no general conceptual basis that combines all the abstractions related to agents.

Representation languages from the theory of agents are used to overcome these methodological difficulties. Such languages are based on some formalism and a programming system that together specify semantics of the language used in programming agents. Though ever newer agent programming languages have been proposed in the literature, only some of them are entirely intelligible from the semantic point of view.

This paper explores the possibility of applying existing formal theories of description of distributed and concurrent systems to interaction protocols for real-time multi-agent systems. In particular it is shown how the language PRALU [Zakrevskij, 1996; Zakrevskij, 1999; Zakrevskij, 2001] proposed for description of parallel logical control algorithms and rooted in the Petri net formalism, can be used to describe agent interaction protocols. The described approach can be used for modeling complex, concurrent conversations between agents in a multi-agent system. It can be used to define protocols for complex conversations composed of a great number of simpler conversations. With the language PRALU it is possible to express graphically concurrent characteristics of a conversation, to capture the state of a complex conversation during runtime, and to reuse described conversation structure for processing multiple concurrent messages. It is demonstrated with a known example of English auction [FIPA, 2002] on how to specify an agent interaction protocol using considered means. Finally, using PRALU language we can verify some key behavioral properties of our protocol description that is facilitated by the use of existing software for the language PRALU [Cheremisinov, 1986; Cheremisinova, 2002; Zakrevskij, 1999].

Then, we introduce a methodology of programming agents in MAS; it is based on the language PRALU. When more complicated, multilayered and concurrent conversations take place among groups of agents, PRALU approach that we use appears to offer advantages over the colored Petri net techniques that are proven to provide the most powerful mechanism for specifying interaction protocols up until now [Bai, 2004]. Further, PRALU as the agent-oriented programming language has such an advantage that its semantics is based on logic formalism, this language allows a simple realization. We show how the behavior of MAS can be simulated entirely in the language PRALU. The offered methodology is based on two-block architecture of organization of the description and realization of MAS that consists of the block of synchronization and the functional block. The first one coordinates performance of parallel processes of agent program, that is, it controls the agent behavior. The second part operates data and carries out calculations connected with complex information structures. The distinctive feature of the methodology is that it allows separating development of the synchronizing part of agent programs from the functional one.

An Agent-Oriented Model

The application domains of MASs are getting more and more complex. Firstly, many current application domains of MASs require agents to work in changing environment (or world) that acts on or is acted on by the system. A closed system is one that has no environment; it is completely self-contained in contrast to an open (uncertain) system, which interacts with its environment. Any real system is open. The MAS must decide what to do and develop a strategy in order to achieve its assigned goals, within open environment. For this, the MAS must have a representation or a model of the environment within which it evolves. The environment is composed of situations. A situation is the complete state of the world at an instant of time.

The application of multi-agent systems to real-time environments can provide new solutions to very complex and restrictive systems such as real-time systems. A suitable method for real-time multi-agent system development must take into account the intrinsic characteristics of systems of this type. As a rule they are distributed, concurrent systems with adaptive and intelligent behavior. For agent-based systems to operate effectively, they must understand messages that have a common ontology underlying them. Understanding messages that refer to ontology can require a considerable amount of reasoning on the part of the agents, and this can affect system performance.

Taking the view of agents as practical systems the predominant approach to specifying agents has involved treating them as intentional systems that may be understood by attributing to them mental states such as beliefs, desires, and intentions [Wooldridge, 1995]. Following this idea, a number of approaches for formally specifying agents have been developed, which are capable of representing the following aspects of an agent-based system:

- The beliefs that agents have – the information they have about their environment, which may be incomplete or incorrect;
- The goals that agents will try to achieve;
- The actions that agents perform and the effects of these actions;
- The ongoing interaction that agents have – how agents interact with each other and their environment over time.

BDI agents are systems that are situated in a changing environment, receive continuous perceptual input, and take actions to affect their environment, all based on their internal mental state. In practical terms, beliefs represent the information an agent has about the state of its environment. It is updated appropriately after each action. The desires denote the objectives to be accomplished, including what priorities are associated with the various objectives. Intentions reflect the actions that must be fulfilled to achieve the goal (the rules to be fired).

Multi-agent conversations are built upon two components:

- Agent communication language;
- Interaction protocol.

There are a number of agent communication languages, such as Knowledge Query and Manipulation Language (KQML) [Finin, 1997] and Agent Communication Language (ACL), proposed by the Foundation for Intelligent Physical Agents (FIPA) [FIPA, 2002], and others designed for special purposes and that are like mentioned ones. These agent communication languages specify a domain specific vocabulary (ontology) and the individual messages that can be exchanged between agents.

Interacting agents should comply with an interaction protocol in order to engage permissible sequences of message exchange. When agent sends a message it can expect a response to be among a set of messages indicated by the accepted protocol. The interaction protocol can be assigned by the designer of the multi-agent system otherwise an agent needs to indicate the protocol that it wants to follow before it starts to interact with other members of the system.

Agent interaction protocols

Interaction protocols [FIPA, 2002] specify the sequences in which messages should be arranged in agent interactions. Protocol constrains number of sequences of allowed messages for each agent at any stage during a communicative interaction, i.e. it describes some standard pattern messages exchanged between agents need to follow. By the very nature of protocols as public conventions, it is desirable to use a formal language to represent them. When agents are involved in interactions where no concurrency is allowed, most conversation protocols are traditionally specified as deterministic finite automata (DFA) [Pitt, 1999] of which there are numerous examples in the literature. DFA consists of a set of states, an input alphabet and a transition function, which maps every pair of states and input to the next state. In the context of interaction protocols, the transitions specify the communicative actions to be used by the various agents involved in a conversation. A protocol based on such a DFA representation determines a class of well-formed conversations. Conversations that are defined in this way have a fixed structure that can be laid down using some kind of graphical representation. However, they are not enough expressive to model complex interactions, especially those with some degree of concurrency.

Protocols can be represented as well in a variety of other ways. The simplest is a message flow diagram, as used by FIPA [FIPA, 2002]. More complex protocols will be better represented using a UML sequence (Unified Modeling Language) [Booch, 1999] and AUML (Agent UML) [Bauer, 2001; Odell, 2001; Winikoff, 2005], interaction diagram, state chart [Harel, 1998] and Colored Petri Net (CPN) [Bai, 2004; Jensen, 1992].

UML is one of the currently most popular graphical design languages that are de facto standard for the description of software systems. AUML extends UML sequence diagrams to support the specification of interactions between agents. The advantage of AUML is its visual representation as well as state chart [Harel, 1998], but unfortunately, AUML suffers from two issues. Firstly, the notation itself is not formally and precisely defined; and secondly, tool support for AUML is largely non-existent [Cost, 1999]. AUML diagrams only offer a semi-formal specification of interactions. This weakness may generate several problems. Indeed, the lack of formal semantics in AUML, such as in UML, can lead to several incoherences in the description of a MAS's behavior. It is difficult, especially in the case of complex MAS, to detect this kind of defects [Poutakidis, 2002].

A CPN model of a system describes the states, which the system may be in, and the transitions between these states. CPNs provide an appropriate mathematical formalism for the description, construction and analysis of distributed and concurrent systems. CPNs can express a great range of interactions in graphical representations and well-defined semantics, and allow formal analysis and transformations [Bai, 2004; Murata, 1989]. By using CPNs, an agent interaction protocol can be modeled as a net of components, which carry the protocol structure. Using CPNs to model agent interaction protocol, the states of an agent interaction are represented by CPN places. Each place has an associated type determining the kind of data that the place may contain. Data exchanges between agents are represented by tokens, and the colors of tokens indicate the data value of the tokens. The interaction policies of a protocol are carried by CPN transitions and their associated arcs. A transition is enabled if all of its input places have tokens, and the colors of these tokens can satisfy constraints that are specified on the arcs. A transition can be fired, which means the actions of this transition can occur, when this transition is enabled. When a transition occurs, it consumes all the input tokens as computing parameters, conducts conversation policy and adds new tokens into all of its output places. After a transition occurs, the state (marking) of a protocol has been changed and a protocol will be in terminal state when there is no enabled or fired transition.

There are a number of works using Petri Nets or CPNs to model agent interaction protocols [Cost, 1999; Poutakidis, 2002] there have been also some works on the investigation of flexibility, robustness and extensibility of protocols [Hutchison, 2002]. Today, only Petri Net models are considered to be one of the best ways to model agent interaction protocols. However the notion of an agent executing an action with Petri Net is not explicit in the notation [Paurobally, 2002]. A different PN can be assigned to each agent role, raising questions about how the entire protocol is inferred and the reach ability and consistency of shared places. If a single Petri net is partitioned for each role, this leads to a complex diagram where a partition is required for each agent identified. Furthermore, alternative actions and states either agree or reject but not both, cannot be expressed in standard Petri nets.

It is necessary for any protocol itself to be correct and verifiable. If it is not correct then the agents that follow it may perform contradictory and unexpected actions leading to possible breakdown of the interaction. The central problem of the verification of interactions that take place in open (not being cooperative) systems is the problem of conformance inspection between behavior of agents and interaction protocol. That is the protocol must be understandable by all agents of the system and they must behave according to this protocol. Thus a language used to represent MAS protocol must be clear enough to allow means for protocol verification. We consider that a language for developing protocols is needed which can ideally meet the following requirements:

1. Provide a graphical representation for ready perception of structure by MAS designer;
2. Have an unambiguous formal specification with clear semantics for verification;
3. Be close to an executable language for implementation purposes;
4. For relative tractability, maintain a propositional form for a formal language;
5. Provide a well-defined program logic for ensuring complete protocols and validating the properties of a protocol;
6. Exhibit enough expressiveness for agent interactions and nested interactions.

Keeping in mind complex systems characterized by complex interaction, asynchronism and concurrency we propose to use for the purposes of their description a special language PRALU [Zakrevskij, 1996; Zakrevskij, 1999; Zakrevskij, 2001] satisfying all mentioned requirements. It has its background in the Petri net theory (expanded nets of free choice – EFC–nets investigated by Hack [Hack, 1972]) but possesses special means for keeping track of the current states of the conversation, receiving messages and initiating responses. The language combines properties of “cause-effect” models with Petri nets. It is intended for a wide application in engineering practice and is well suited for representation of the interactions involved in concurrent system; synchronization among them and then it is simple enough for understanding. The language PRALU supports hierarchical description of the algorithms that is especially important in the case of complex systems. At last powerful software has been developed that provides correctness verifying, simulation, hardware and software implementation of PRALU-descriptions [Cheremisinov, 1986; Cheremisinova, 2002]. The review of obtained results in this field one can find in [Cheremisinova, 2002].

PRALU language

Any algorithm in PRALU consists of sequences of operations to be executed in some pre-determined order. Two basic operations are used in PRALU: acting “$\rightarrow A$” and waiting “$- p$” operations. The first one changes the state of the object under control, whereas the second one is passive waiting for some event without affecting anything. In simple case A and p are conjunctive terms, so acting and waiting operations can be interpreted as waiting for event p = 1 and producing event A = 1. But A can be understood too as a formulae defining operations to be performed and p as a predicate defining condition to be verified. For example, acting and waiting operations could be specified by the expressions such as

$$-(a > b + c) \text{ and } \rightarrow (a := b + c).$$

The sequences consisting of action and waiting operations are considered to be linear algorithms. For instance, the following expression means: wait for p and execute A, execute B, then wait for q and execute C:

$$- p \rightarrow A \rightarrow B - q \rightarrow C.$$

In general, a control algorithm can be presented as an unordered set of chains α_j in the form

$$\mu_j : - p_j \, L_j \rightarrow \nu_j,$$

where L_j is a linear algorithm, μ_j and ν_j denote the initial and the terminal chain labels being some subsets of integers from the set $M = \{1, 2, ..., m\}$: $\mu_j, \nu_j \subset M$ and the expression “$\rightarrow \nu_i$“ presents the transition operation: to the chains with labels from ν_j.

Chains can be fulfilled both serially and in parallel. The order in which they should be fulfilled is determined by the variable starting set $N_t \subseteq M$ (its initial value N_0 = {1} as a rule): a chain $\alpha_j = \mu_j : - p_j L_j \rightarrow \nu_j$ (that was passive) is activated if $\mu_j \subseteq N_t$ and p_j = 1. After executing the operations of the algorithm L_j, N_t gets a new value $N_{t+1} = (N_t \setminus \mu_j) \cup \nu_j$. The algorithm can finish when some terminal value of N is reached (one-element as a rule), at which time all chains became passive. But the algorithms can also be cyclic; they are widely used when describing production processes.

When the conditions $\mu_j \subseteq N_t$ ($|\mu_j| > 1$) and p_j = 1 are satisfied for several chains simultaneously these chains will be fulfilled concurrently. On the contrary chains with the same initial labels are alternative (only one of them can be fulfilled at a time), they are united in a sentence with the same label as will be shown below.

Thus PRALU allows concurrent and alternative branching, as well as merging concurrent and converging alternative branches. These possibilities are illustrated with the following examples of simple fragments [Zakrevskij, 2000]:

Concurrent branching	Merging concurrent branching	Alternative branching	Converging alternative branching
1: ...$\rightarrow$2.3	2: ...$\rightarrow$4	1: $- a$...$\rightarrow$2	2: ...$\rightarrow$4
2: ...	3: ...$\rightarrow$5	$- \bar{a}$ ...$\rightarrow$3	3: ...$\rightarrow$4
3: …	4.5: ...		4: ...

There exist in PRALU two syntactic constraints on chains that restrict concurrent and alternative brunching. If some chains are united in the same sentence (they have the same initial labels) they should have orthogonal predicates in the waiting operations opening the chains:

$$(i \neq j) \ \& \ (\mu_i \cap \mu_j \neq \varnothing) \rightarrow (p_i \ \& \ p_j = 0).$$

The other constraint is similar to the corresponding condition specific for extended nets of free choice (Hack [Hack, 1972]):

$$(i \neq j) \ \& \ (\mu_i \cap \mu_j \neq \varnothing) \rightarrow (\mu_i = \mu_j).$$

PRALU language has some more useful properties that can be useful for description of complex interaction protocols.

1. PRALU algorithms can be expressed both in graphical and symbolic forms.
2. PRALU language permits hierarchical description of protocols. The two-terminal algorithms (having the only initial and the only terminal chain labels) may be used as blocks (invoked as complex acting operations) in hierarchical algorithms.
3. In PRALU there exist some additional interesting operations that can be useful when describing interaction protocol. Among those are suppression operations that provide response on special events that can take place outside or within control systems. Suppression operations

   “$\rightarrow*$”, “$\rightarrow*\gamma$”, “$\rightarrow'\gamma$” (where $\gamma \subseteq M$)

   interrupt the execution of all concurrently executed algorithm chains (in the case of “$\rightarrow*$”), only those ones mentioned in γ (in the case of “$\rightarrow*\gamma$”) or are not mentioned in γ (in the case of the operation “$\rightarrow'\gamma$”). These operations break the normal algorithm flow.
4. In addition to logical variables, arithmetic variables are also used in the PRALU language. In particular, among arithmetic operations ought to be mentioned the following operations are introduced:
   – Timeout operation “– *n*” – delay for *n* unit times;
   – Counting operations that count event occurrences:

     “(*x* = *n*)” – assignment of a natural value *n* to a multi-valued variable *x*;

     “(*x* +)” and “(*x* –)” – assignment of unit positive and unit negative increment in value;

     “–(*x* = *n*)” – awaiting the start of an event: the value of *x* is equal to *n.*

Representing Agent Interaction Protocols in PRALU

As an example of an interaction protocol let consider English auction [FIPA, 2002]. The auctioneer seeks to find the market price of a good by initially proposing a price below that of the supposed market value and then gradually raising the price. Each time the price is announced, the auctioneer waits to see if any buyers will signal their willingness to pay the proposed price. As soon as one buyer indicates that it will accept the price, the auctioneer issues a new call for bids with an incremented price and continues until no buyers are prepared to pay the proposed price. If the last price that was accepted by a buyer exceeds the auctioneer's (privately known) reservation price, the good is sold to that buyer for the agreed price. If the last accepted price is less than the reservation price, the good is not sold.

In the case of the auction there are participants of two types: the Auctioneer and Buyers. So, we have two kinds of interaction protocols – those of Auctioneer and of Buyers. So, we have two kinds of interaction protocols – those of Auctioneer and of Buyers. The last participants are peer and should be described with identical interaction protocols.

Interaction protocol as a whole can be represented in PRALU as three complex acting operations – blocks that are exchanging with values of logical variables, only such variables are mentioned in them. Each block has some sets of inputs and outputs that are enumerated in brackets following the block name (the other variables of a block are its internal). Initialization of a block algorithm is depicted by the fragment such as "→∗Buyer". The operation Buyer exists in as many copies as the number of participants of the auction, the copies differ in their indexes only.

The modeling of the process of auction begins with the execution of "*Main_process*" triggering event that initiates the interaction protocol execution. Here the processes Auctioneer and *Buyer$_n$s* are executed concurrently. For the sake of simplicity we limit the number of buyers to two. The process Auctioneer starts with sending the first message (start_auction) that is waited by others participants to continue communication.

Below PRALU description of the auction interaction protocol is shown. Here we show the only block *Buyer$_n$*, but for real application (intending to simulate the process of auction, for example) we should have as many proper copies as it has been used (in our case – two). Here through "′var1" the inversion of the Boolean variable var1 is denoted.

Main_process ()

1: → 2.3.4

2: →∗Auctioneer → 5

3: →∗Buyer$_1$ → 6

4: →∗Buyer$_2$ → 7

5.6.7: → .

Buyer$_n$ *(start_auction, price_proposed, end_auction / accept_price$_n$, not_understand)*

1: – start_auction → 2

2: – price_proposed →∗Decide (/ decision_accept, decision_reject) → 3

– end_auction → **.**

3: – decision_accept → accept_price$_n$ → 4

– decision_reject → ′accept_price$_n$ → 4

– not_understand → 4

4: – timeout → 2

Auctioneer *(accept_price$_1$, accept_price$_2$, not_understand / start_auction, price_proposed, end_auction)*

1: → start_auction → 2

2: → ′accept_price$_1$.′accept_price$_1$.′not_understand →∗Price_propose (/price_proposed) → 3

3: – not_understand → 2

– accept_ price$_1$ → 4

– accept_ price$_2$ → 4

–′not_understand.′accept_ price$_1$.′accept_ price$_2$ → end_auction →

∗Is_reservation_price_exceeded (/ is_exceeded) → 6

4: →′price_proposed.′win$_1$.′win$_2$ → 5

5: – accept_ price$_1$ →win1 → 2

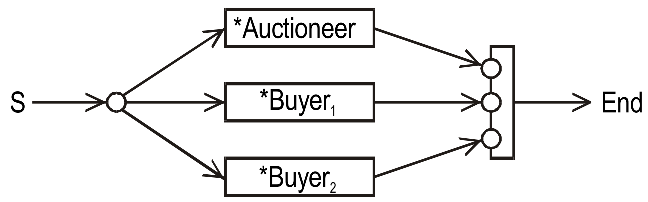
– accept_ $price_2$ →win2 → 2

6: – is_exceeded →good_sold → 7

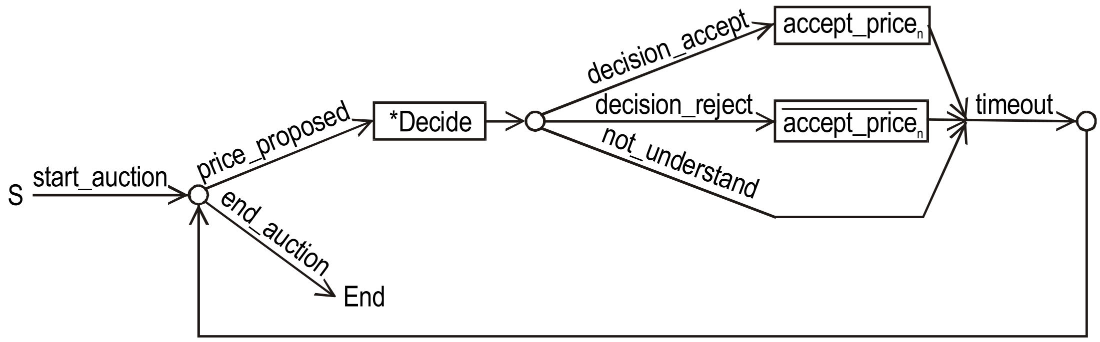
– ′is_exceeded →′good_sold → 7

7: → .

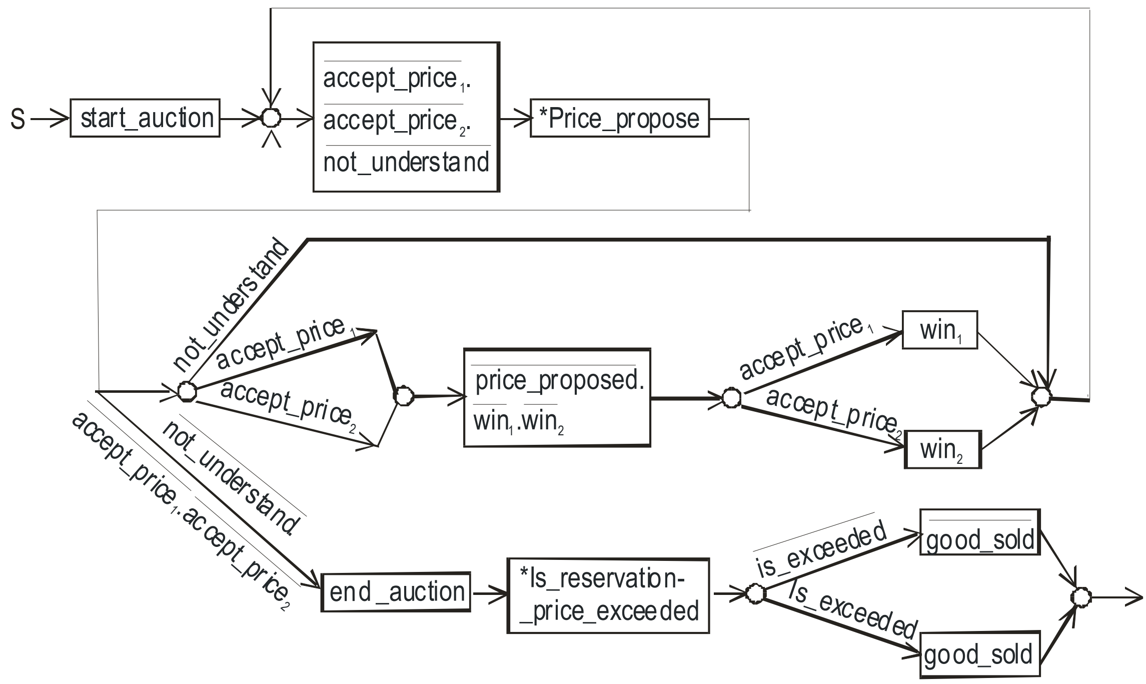
Figure 1 depicts graphical schemes of three mentioned PRALU-blocks: Main_process, Auctioneer and $Buyer_n$.



Main_process ()



$Buyer_n$ (start_auction, price_proposed, end_auction / $accept\_price_n$, not_understand)



Auctioneer ($accept\_price_1$, $accept\_price_2$, not_understand / start_auction, price_proposed, end_auction)

Figure1. English auction interaction protocol in PRALU

It is assumed that all unformalized operations are referred to as acting operations that set values of logical variables assigned to them. For example, Buyer’s operation “Decide” decides for accepting or rejecting the announced price. Depending on adopted decision, it outputs true value of logical variable “decision_accept” or of logical variable “decision_reject”. In a similar, Auctioneers operation “Price_propose” proposes an initial price or increments the charged price outputting true value of logical variable “price_proposed”; the operation “Is_reservation_price_exceeded” verifies if the price accepted by a buyer exceeds the auctioneer’s reservation price outputting true or false value of logical variable “is_exceeded”.

The operation “– timeout” (where “timeout” is an integer number) means waiting for “timeout” unit times before doing something followed it. The operation “→.” is interpreted as the transition to an end of the process described by the block. When the processes of Auctioneer and all of Buyers reach their end in the Main_process the transition to its finish is executed.

BDI architecture for the language PRALU

In practical programming, an agent is a well-formed entity incorporated in a computer system designed to perform flexible, independent actions for the purpose of attaining specified goals. Agents differ from ordinary software by the complexity of the interaction and communication scenarios. Any physical multi-agent system is open, i.e., the agents function in a varying environment, which affects the agents and varies as a consequence of their operations. The system must arrive at decisions so as to attain the objectives assigned to it, and for this purpose it must possess a model of the environment in which its behavior evolves.

Over the last decade, the specification and application of BDI agents (belief, desire, intention) [Burmeister, 1992] have received a great deal of attention. Within the BDI architecture agents are associated with beliefs (typically about the environment and other agents), desires or goals to achieve, and intentions or plans to act upon to achieve its desires. These three components completely set a state of the “mind” of the agent.

In programming terms, beliefs represent knowledge (information) a BDI agent has about the state of the environment; that is updated after each action. The desires denote the objectives to be accomplished, taking into account their priorities. Intentions reflect the actions that should be fulfilled to achieve the goal (the rules to be fired). Interaction protocols reduce search space of possible decisions, owing to limited range of answers to possible messages for the given situation. Ontology [Gruber, 1993] as a formal specification of terms of a subject domain and relations between them is almost equivalent to the notion of semantics of programming language.

In our case ontology will define the terms of BDI architecture in terms of the language PRALU. It is possible to consider, that BDI agent consists of the sets of beliefs *B*, plans *P*, situations *E*, actions *A* and intentions *I*. When the agent notes a change in its environment, it decides, that there an event takes place representing some situation from *E*. Registration of the event consists in changing a state of the agents “mind”: a choice of some belief from *B*. According to it and the desire (defined by some plan from *P*) the agent intends to execute some intention representing some sequence of actions from *A* to achieve the goal chosen from *I*. Thus, planned actions are defined by the chosen plan from *P*. After they were carried out, the current situation of the environment is changed.

In traditional parallel programming languages the basic concepts are data and calculation control. Data are represented by values of variables, and the control is specified by a set of processes which transform local memory states defining variable values. The concept of the intellectual agent introduces new ideas of data manipulation in parallel programming. Agent’s states are set by more complex information structures of predicates logic of the first order or modal logic. Calculations as transformations of memory states are controlled by protocols specifying communications between agents.

The program of a BDI agent is a set of plans defining actions by means of which the agent should reach a goal of its functioning. The plan consists of head labels, a body and tail labels. The body of the plan is a sequence of actions, by means of which the agent should reach a goal of its functioning, and conditions which the agent should check up. The head and tail labels of the plan symbolize intentions. Critical concept for the behaviour of the agent is the concept of active intention. The plan will be carried out only when all intentions from its head are active. After executing the plan all intentions from the head become inactive and intentions from the tail quite the contrary become active. The current set of active intentions always is not empty, the body of a plan and a set of tail intentions can be empty. Such a model of BDI agent can be easily described on PRALU: a plan has a direct analogy with a chain α_j (in the form $\mu_j : - p_j L_j \to \nu_j$, L_j: “$-p \to A -q \to B -r \to C\ldots$”). The action leading to some goal can be described by acting operation “$\to g$” (done g) that may be executed, if some belief has appeared correct. Check of such condition is described by waiting operation “$-$ a” (happens a). An agent carries out the action following waiting operation only after the moment when “belief a” becomes true.

Methodology of programming agents on PRALU

We propose the natural methodology of designing agent program based on splitting the program into two parts: synchronization and functional blocks (Figure 2). The first block coordinates performance of parallel processes of the agent program, that is, it controls the agent behavior. The synchronization block should be described on PRALU language. The functional block operates data and carries out calculations connected with complex information structures (formulas of predicate or modal logic). This block is realized in programming language.

Such a splitting is carried out at the level of the project statement of MAS. The functional part is presented by predicates that describe memory states of the agent program (or/and external environment) or prescribe performance of some actions. In PRALU-description the appropriate logical variable is introduced for each predicate, setting true value to this variable starts the process of the predicate calculation. At the stage of verification of logical consistency it is better to interpret predicates as independent logical variables.

For example, the predicate *Decide* is used in the description of the behavior of agent Buyer. In the synchronization block in Figure 3 it is represented by the Boolean variable *Decide*. In the functional block, the meaning of a predicate is indicated by pseudo-code, which shows the use of the variable *Decide* and rules for computing two other Boolean variables, *decision_accept* and *decision_reject*.

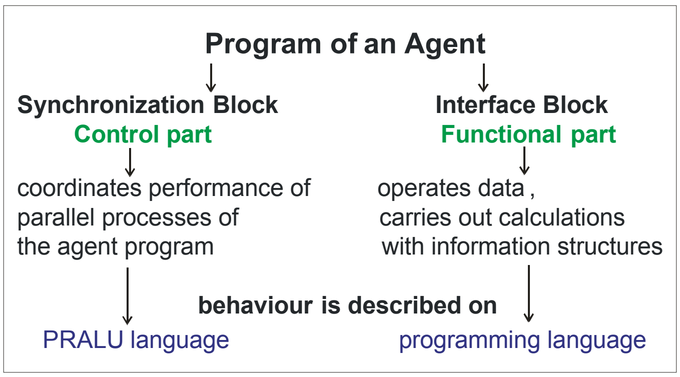


Figure 2. Splitting an agent program

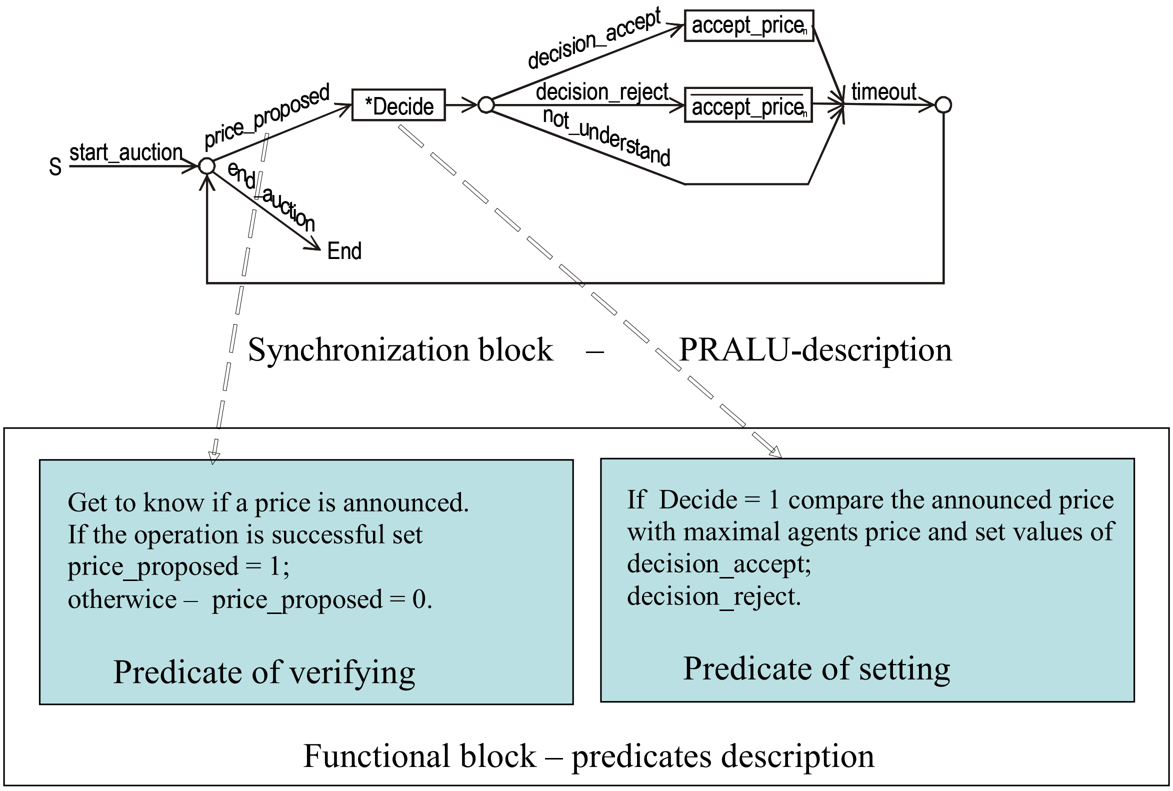


Figure 3. The program of the agent Buyer in English auction

Such an approach allows separating development of the synchronization part of the PRALU-description from the functional one. When designing MAS, the implementation of the functional part can be delayed concentrating designer's attention on working out the synchronization block implementing interaction protocol. This may significantly simplify the process of verifying the logical consistency of the behaviors of agents.

Program implementation of PRALU-descriptions

Programming in PRALU, the designer of the agent specifies sets of its belief, plans and intentions. The process of designing an agent program consists of the following stages.

1. Splitting at functional level the specification of MAS to be designed into synchronization and functional parts.
2. Development and analysis of PRALU-description of the synchronization part of MAS, being based on informal specification of its interaction protocol.
3. Verification of logic consistency of the MAS behavior by checking correctness and simulation of PRALU-description. Elimination of mistakes found out.
4. Program implementation of the functional part of MAS.
5. Development of computer program implementing behavior of MAS: translating PRALU - description of MAS on programming language and binding it with the programs of the functional part.
6. Testing of the generated programs.

The use of the PRALU language as a tool for representing the specifications of agents as well as the executability of these specifications makes the process of designing a multiagent system a structured process. All formal tasks may be implemented in the language and software tools for the implementation of these tasks are available. Moreover, there is the powerful theory supported with software tools for verification, simulation, hardware and software implementation of the descriptions of algorithm entity incorporated in a computer system in the PRALU language [Cheremisinov, 1986; Cheremisinova, 2002; Zakrevskij, 1999].

In the process of program implementation of PRALU-description it is translated on the intermediate language applied in the simulation program [Cheremisinov, 1986] too. The program that is generated by the PRALU compiler consists of two interacting blocks – 1) calculation of responses, and 2) control of sensors and servo mechanisms of the system. The control structure of the response computation program consist in an infinite cycle involving input of signals from the agent's sensors that constitute its beliefs into the internal memory, computation of the agent's responses from the values of these signals, and output of signals to the agent's servo mechanisms. All communication related to control and to data between the response computation block, sensor control block, and the servo mechanisms are input into the representation of the PRALU description implemented in the intermediate language.

To implement a parallel algorithm on the only processor [Cheremisinov, 1986], it is necessary to order the operations of the intermediate language by means of the PRALU intermediate language planner, which has its own properties and methods. The properties of the planner consist in the presence of two queues, a queue of waiting and a queue of ready branches. The methods of a program planner consist in triggering, halting, and pausing subprocesses. In the planner's initial state the queue of ready branches contains the first operation of the algorithm while the queue of waiting branches is empty. The program automatically orders the sequence of execution of operations that are in the process of being executed and there is no need for preliminary planning.

In each cycle the planner extracts the next branch from the queue of ready branches, executes a corresponding chain of algorithm in PRALU, and places all the branches that must be executed following the given branch into the waiting queue. If the planner discovers that the queue of ready branches is empty, (1) it transfers the contents of the queue of waiting branches to the queue of ready branches, rendering the queue of waiting branches empty; and (2) executes information exchange with the environment. Such a sequence of execution of chains of an algorithm in PRALU guarantees that operations of the initial algorithm that may be executed in parallel are all of the same length. Thus, the program implementation of PRALU possesses the semantics of measurable time that satisfy a rendezvous condition [Cheremisinov, 2008].

In the two-block model of the program, the predicates of the functional part are considered as a "supplementary tool" of sensors and servo mechanisms of the agent that generates logical signals or is triggered by a signal. If the results of the execution of computational operations must be taken into account in the control process, the supplementary tool will generate logical signals expressly for this goal. In the program that controls the sensors and servo mechanisms, this supplementary tool is described in the form of expressions in the ordinary programming language by the designer of the agent's program.

This strategy of programming of agents is especially simple where the Forth programming language is used as a platform [Moore, 1974]. In this case it is only necessary to define a single word of Forth for each operator of an intermediate language.

The principle of implementation of the Forth language, i.e., threaded code, may be used for programming of agents in a C (or C++) language platform. The threaded code, a strategy used in the Forth language, consists in representation of a program in which nearly all elements are denoted by procedure calls (some procedures presuppose that data are located behind them in the threaded code). A stack is used to implement procedure calls; to distinguish it from a data stack, such a stack is referred to as a return stack and is used for storage of procedure data. Queues of ready and waiting branches of the intermediate language are implemented by means of return and data stacks, respectively.

The transformation of a representation of an algorithm in PRALU into a C program is executed by a single-pass text translator that converts the operators of the intermediate language into procedure calls in C language syntax. The efficiency of a program constructed in this way may be estimated using as an example the implementation of one of the most difficult operations of the intermediate language, the pause operation. The job executed by this operation consists in storage of the address of the next procedure call in the data stack and a transfer to execution of a procedure the address of which is extracted from the return stack. The C compiler constructs a fragment of code consisting of only two machine instructions for execution of these operations.

When programming agents on the base of a C language platform, procedures defined by predicates should be written in C using the descriptions like in Figure 3 as a model.

Most of the well-known systems of logical programming of agents use a computation model from the Prolog language [Endriss, 2004]. Prolog is based on first-order predicate logic and the objects which the program manipulates are symbols without any meaningful interpretation. Execution of a Prolog program involves a process called unification and consists in scanning a database of facts and selecting values that satisfy some query. Unification is based on syntactic identity. In order to find a set of solutions, Prolog traverses a search tree, tests a set of variants, most of which do not occur in a solution, and, if unsuccessful, returns to a previous state and tests a different branch. For complex problems this search process requires great expenditure of memory and time, which is one reason for the computational inefficiency of Prolog.

In comparing the proposed PRALU-based logical programming system with other systems, for example, Prolog-based systems [Endriss, 2004], the improved computational efficiency of PRALU-based systems is apparent; in fact, the computational efficiency of PRALU-based systems is comparable with that of C language programs.

Software tools for automatic development, debugging, hardware and software implementation have been developed for the PRALU language [Cheremisinov, 1986; Cheremisinov1, 1986; Cheremisinova, 2002]. In particular, a modeling program may be used to display the behavior of new protocols written in PRALU. Display of behavior consists in exhibiting the control states and states of the variables of an algorithm in PRALU. The control states are exhibited by distinguishing operations in the text of the algorithm that are being executed at a given time (display of activity points), as it is done in most debuggers.

Conclusion

This paper has addressed the need for formalized and more expressive logical and graphical methodologies for specifying (and then validating) interaction protocols in multi-agent systems. Towards this, it was proposed to use the formal language PRALU intended for the representation of complex interactions involved in concurrent system, being in need of synchronization among these interactions. It was demonstrated as well how PRALU algorithms could be used for the specification of multi-agent interaction protocols by the example of English auction. In favor of using the language PRALU is the existence of a great deal of methods and software developed for simulation of PRALU algorithms as well as for their hardware and software implementation.

The methodology is suggested that ensures structuring the process of MAS designing on the base of PRALU language by means of separating calculation part of MAS specification from its control part. That allows the designer to concentrate on more complex stage of MAS designing – its interaction protocol. It is shown that language PRALU is very suitable for specifying interaction protocols of MAS and for implementation of suggested methodology. The obtained results are directed towards the usage when designing parallel/distributed technical systems too.

Acknowledgement

The paper is published with partial support by the project ITHEA XXI of the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

Bibliography

[ARPA, 1993] Specification of the KQML Agent-Communication Language, ARPA Knowledge Sharing Initiative, External Interfaces Working Group, July 1993

[Bai, 2004] Quan Bai, Minjie Zhang Khin, Than Win. A Coloured Petri Net Based Approach for Multi-agent Interactions. In: The 2nd International Conference on Autonomous Robots and Agents, Palmerston North, New Zealand, 13 – 15 December 2004.

[Bauer, 2001] B. Bauer, J.P. Müller, J. Odell. Agent UML: A Formalism for Specifying Multiagent Interaction. In: Agent-Oriented Software Engineering, Ed. P. Ciancarini and M. Wooldridge, Springer-Verlag, Berlin, 2001.

[Booch, 1999] G Booch, J. Rumbaugh, I. Jacobson. The Unified Modelling Language User Guide, Addison Wesley, 1999.

[Burmeister, 1992] B. Burmeister, K. Sundermeyer. Cooperative problem-solving guided by intensions and perception. In: Demazean Decentralized A.I. 3. Ed. E. Werner and Y. Amsterdam, The Netherlands, North Holland, 1992.

[Cheremisinov, 1986] D.I. Cheremisinov. Implementation of parallel algorithms on a microprocessor. In: Programming, 1986, No 1 (in Russian).

[Cheremisinov, 2008] D. Cheremisinov, L. Cheremisinova. Formalization of interaction events in Multi-agent Systems. In: Information Technologies & Knowledge (IJ ITK), 2008, Vol. 2, No. 2.

[Cheremisinov1, 1986] D.I. Cheremisinov. LyaPAS-M Programming System Based Package for the Development of Microprocessor Programs. In: Upravlyayushchie sistemy i mashiny, 1986, No 1 (in Russian).

[Cheremisinova, 2002] L.D. Cheremisinova. Realization of parallel algorithms for logical control. Institute of Engineering Cybernetics of NAS of Belarus, Minsk, 2002 (in Russian)

[Cost, 1999] R. Cost. Modelling Agent Conversations with Coloured Petri Nets. In: Working Notes of the Workshop on Specifying and Implementing Conversation Policies, Seattle,Washington, 1999.

[Endriss, 2004] U. Endriss, N. Maude, F. Sadri, F. Toni. Logic-Based Agent Communication Protocols. In: Advances in Agent Communication, Ed. F. Dignum, Laboratory Notes in Artificial Intelligence, 2004, Vol. 2922.

[Finin, 1997] F. Finin, Y. Labrou, J. Mayfield. KQML as an agent communication language. In: Software Agents, Ed. J.M. Bradshaw, MIT Press, 1997.

[FIPA, 2002] Foundation for Intelligent Physical Agents (FIPA), Communicative Ac Library Specification, 2002; http://www.fipa.org/specs/fipa00037.

[Gruber, 1993] T.R. Gruber. A Translation Approach to Portable Ontologies. In: Knowledge Acquisition, 1993, Vol. 5, No 2.

[Hack, 1972] M. Hack. Analysis of production schemata by Petri nets. Project MAC-94, Cambridge, 1972.

[Harel, 1998] D. Harel, M. Politi. Modelling reactive systems with statecharts. McGraw-Hill, 1998.

[Hutchison, 2002] J. Hutchison, M. Winikoff. Flexibility and Robustness in Agent Interaction Protocols. In: Proceedings of the 1st International Workshop on Challenges in Open Agent Systems, Bologna, Italy, 2002.

[Jennings, 2001] N. R. Jennings. An Agent-Based Approach for Building Complex Software Systems. In: Communications of the ACM, 2001, Vol. 44, No 4.

[Jensen, 1992] K. Jensen. Colored Petri Nets – Basic Concepts. In: Analysis Methods and Practical Use, Springer-Verlag, Berlin, 1992, Vol. 1.

[Lesser, 1999] V. Lesser. Cooperative Multiagent Systems: A Personal View of the State of the Art. In: IEEE Trans. on Knowledge and Data Engineering, 1999, Vol. 11, No 1.

[Moore, 1974] C.H. Moore. FORTH: A New Way to Program a Mini-computer. In: Astro. Astrophys. Suppl., 1974, Vol. 5.

[Murata, 1989] T. Murata. Petri Nets: Properties, Analysis and Applications. In: Proceedings of the IEEE, 1989, Vol. 77, No 4.

[Odell, 2001] J.J. Odell, H.V.D. Parunak, B. Bauer. Representing Agent Interaction Protocols in UML. In: Agent-Oriented Software Engineering, Ed. P. Ciancarini and M. Wooldridge, Springer-Verlag, Berlin, 2001.

[Paurobally, 2002] S. Paurobally, J. Cunningham. Achieving Common Interaction Protocols in Open Agent Environments. In: Proceedings of 1st International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS'02), Bologna, Italy, July 2002.

[Pitt, 1999] J. Pitt, A. Mamdani. Protocol-based Semantics for an Agent Communication Language. In: Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-1999), Morgan Kaufmann Publishers Inc., San Francisco, USA, 1999.

[Poutakidis, 2002] D. Poutakidis, L. Padgham, and M. Winikoff. Debugging Multi-Agent Systems Using Design Artefacts: the Case of Interaction Protocols. In: Proceedings of the 1st International Joint Conference on Autonomous Agents and Multi Agent Systems, Bologna, Italy, 2002.

[Searle, 1986] J.R. Searle. What is a Speech Act. In.: Novoe v zarubezhnoy lingvistike, 1986, issue 17 (in Russian).

[Shannon, 1948] C.E. Shannon. A Mathematical Theory of Communication. In: Bell Syst. Tech. J., 1948, Vol. 27.

[Subrahmanian, 2000] V.S. Subrahmanian, P. Bonatti, J. Dix et al. Heterogeneous Agent Systems. MIT Press, 2000.

[Winikoff, 2005] M. Winikoff. Towards Making Agent UML Practical: A Textual Notation and a Tool. In: Proceedings of the Fifth International Conference on Quality Software, September 19–20 2005,.

[Wooldridge, 1995] M. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. In: The Knowledge Engineering Review, 1995, 10(2).

[Zakrevskij, 1996] A.D. Zakrevskij. Parallel logical control algorithms: verification and hardware implementation. In: Computer Science Journal of Moldova, 1996, Vol. 4, No 1.

[Zakrevskij, 1999] A.D. Zakrevskij. Parallel algorithms for logical control. Minsk, Institute of Engineering Cybernetics of NAS of Belarus, 1999 (in Russian).

[Zakrevskij, 2000] A. Zakrevskij, V. Sklyarov. The Specification and Design of Parallel Logical Control Devices. In: Proceedings of PDPTA'2000, June, Las Vegas, USA, 2000.

[Zakrevskij, 2001] A. Zakrevskij, L. Zakrevski. Algorithms for logical control: their description, verification and hardware implementation. In: Proceedings of PDPTA'01, Las Vegas, 2001, Vol. 2.

Authors' Information

Dmitry Cheremisinov – The United Institute of Informatics Problems of National Academy of Sciences of Belarus, leading researcher, Surganov str., 6, Minsk, 220012, Belarus; e-mail: cher@newman.bas-net.by

Major Fields of Scientific Research: Logic design automation, System programming

Liudmila Cheremisinova – The United Institute of Informatics Problems of National Academy of Sciences of Belarus, principal researcher, Surganov str., 6, Minsk, 220012, Belarus; e-mail: cld@newman.bas-net.by

Major Fields of Scientific Research: Discrete mathematics, Logic design automation

A LANGUAGE USING QUANTIFIERS FOR DESCRIPTION OF ASSERTIONS ABOUT SOME NUMBER TOTAL PASCAL FUNCTIONS

Nikolay Kosovskii

***Abstract**: Mathematical models of N-finite ($N \geq 2$) (especially for N= 2^{16} typical for IBM-compatible personal computers) real, integer or Boolean valued computations are proposed in the frameworks of a finite-discrete mathematics. Such a mathematical model more precisely characterizes the contemporary practice of the computer use. Notions of an N-finite RAM (Random Access Machine), an N-finite RASP (Random Access Machine with Stored Program) and an N-finite Pascal program using integer operations modulo N with the reminders from the segment $[-\lfloor N/2 \rfloor, \lfloor N/2 \rfloor - 1 + (N \bmod 2)]$ are introduced. A formal N-finite quantifier logic language of modulo N arithmetic operations is defined. It allows formulating mathematical properties of such a non-recursive total Pascal function or predicate. It is proved that the problem of a formula identical truth in such a language containing arithmetic operations modulo N is a P-SPACE-complete one. So, the proof of a correctness of a non-recursive total Pascal program may be simplified with the use of the proposed language.*

***Keywords:** P-SPACE-completeness, RAM, RASP, Pascal, correctness of a Pascal program.*

***ACM Classification Keywords:** F.2.m Analysis of Algorithms and Problem Complexity Miscellaneous.*

Introduction

Every computer program is a mathematically precise one. But some assertions about a program run result may be rather ambiguous. A language for the mathematically precise notation of properties, conditions and assertions about some Pascal programs is proposed below.

A Pascal function is a total one if it halts and has a value for any input data. For example, a Pascal function is a total one if the following conditions are fulfilled:

- Every division of two numbers is a total one (i.e. before performing the division there is an auxiliary checking whether the divisor differs from zero and if the divisor is zero then the program gives some value for the result);
- The operator *goto* backward is forbidden;
- The loop header begins only with *for*;
- Every non-empty set of Pascal functions (even a set of one function) must not be mutually recursive one;
- Input statements and files are not used.

Let's consider a Pascal function as a total one. Even with such a restriction the first order logic language signature is not sufficient for formal description of properties of an array with integer indexes.

Let a Pascal program uses variables only of the types *real*, *integer, Boolean* and arrays of these types. The set of all numbers of these types usually contains only a finite amount of distinct numbers. For example, a maximal amount of distinct numbers of the types *real* and *integer* for a personal computer is 2^{16} or 2^{32} or sometimes 2^{64}. This illustrates that the use of the set of all integers as a possible set of input data is evidently excessive. Besides, it allows (as it is shown below) essentially to simplify the checking of identical truth for assertions about a Pascal program formulated by means of the proposed extension of the first order logic language.

Besides, a practical number computation mostly uses not integers with unbounded length of their notations but only integers from $[-2^{15}, 2^{15} - 1]$ (the computer type *integer*) or integers from $[-2^{31}, 2^{31} - 1]$ (the computer type *long integer*). And all IBM-compatible computer computations are fulfilled modulo 2^{16} or 2^{32} respectively. That is why traditional mathematical models of computation are not useful for detailed mathematical simulation of a practical personal computer number program.

To formulate mathematical properties of a program it is convenient to use the formal arithmetic language with additional total functions and total predicates of some finite signature. It is possible because every function and every predicate is a total one. At the same time the totality of a Pascal function is an important aspect of friendly software.

That is why the condition of a program totality is natural for the contemporary programmers. In particular, assertions about the total correctness of a program based on the Hoare triads [5, 7] for a Pascal total program may be formulated in such a language.

Below a notion of *N*-finite data of the types *real* and *integer* as well as a notion of the first order *N*-finite index formula are introduced to prove a P-SPACE-completeness of the problem of an *N*-finite index first order formula identical truth. This is impossible if an algorithm (a program) is used to process data with arbitrarily large integers, in particular, because of the use in the formulas quantifiers upon the set of all integers.

Hence, the proposed below mathematical models of computation and formal language for a condition or an assertion setting about the result of such a computation is more adequate to the contemporary practical computation by a program (for example, by a Pascal program) than the use of arbitrarily large numbers as input data.

Initial definitions

A number of the type *integer* from the segment $[-\lfloor N/2 \rfloor, \lfloor N/2 \rfloor - 1 + (N \bmod 2)]$ is called here an *N-integer*. Here $\lfloor N/2 \rfloor$ denotes the maximal integer not greater than N/2 and (N mod 2) denotes the remainder of *N* divided by 2.

A floating-point representation of a number of the type *real* with the notation length not greater then $\lfloor \log_2 N \rfloor$ bits is called here an *N-real* one. The amount of such *N-real* numbers is not greater than *N*.

Pascal functions and Pascal predicates dealing only with numbers and variables of the types *N-real* and *N-integer* are called here *N*-finite ones.

In such a case the use of a traditional for the Pascal language [3] notation for the call of a Pascal function or a Pascal predicate inside an expression (below the term "*N*-finite index term in a finite signature" is used) does not require a more detailed definition than that of a Pascal expression computation. In particular, if an expression contains an *N-real* number then the value of this expression has the type *N-real*.

The notions of RAM (Random Access Machine) and RASP (Random Access Machine with Stored Program) are defined in [1] as models for a non-negative integer valued computation using the only one (actually dynamical) linear array for integers with unbounded notation lengths.

Let's define the notions of an ***N*-finite RAM** and an ***N*-finite RASP** using arithmetic operations modulo *N* with the reminders from the segment [$-\lfloor N/2 \rfloor, \lfloor N/2 \rfloor - 1 + (N \bmod 2)$].

It is suggested to use the only one array of integers from this segment with arbitrary number of dimensions (one-dimensional, two-dimensional, ...) inside a program.

The number of elements in every dimension does not exceed *N*. The input data for an *N*-finite RAM and an *N*-finite RASP are situated in the several first dimensions but the *N*-finite RASP program (with the counter of commands) is situated in the last dimensions. Besides, the counter of commands is situated in the very last dimensions. The bound of an additional dimension uses the only one constant (for example, 1). This additional dimension may be used for the indication of the end of the input data (and the beginning of an *N*-finite RASP program as well as the beginning of the command counters).

So, the array for every *N*-finite RAM or *N*-finite RASP program contains the only finite number of elements. More exactly, this number must be bounded by N, N^2, ..., because of potentially increasing number of array dimensions both for *N*-finite RAM and *N*-finite RASP programs. The last models are the base for interpretation of an arbitrary *N*-finite Pascal function.

The notion of a formula of the first order many sorted logic language is not sufficient if we use a traditional notation for an array element in a Pascal language expression. That is why the notion of such a logic formula with the calls of *N*-finite number Pascal total functions and *N*-finite number Pascal total predicates of a finite signature is extended up to the notion of an *N*-finite index formula of a finite signature.

First of all, the notion of an ***N*-finite index term of a finite signature** is introduced. Its definition may be received from the definition of the *term of a finite signature* (see, for example, [6, 7]) by means of replacing the word "term" by the words "*N*-finite index term of a finite signature" and adding to the definition the following construction: a word of the form “array variable followed by an included into the square brackets sequence of *N*-finite index terms of a finite signature” is an *N*-finite index term of a finite signature. Every *N*-finite index term of this sequence must have the same enumerated type as it was defined in the description of the used array. Only *N*-finite Pascal total functions and arithmetic total operations with *N-reals* and *N-integers* as well as logical constants and operations are used in the new definition. As it is usual for the Pascal language an infix notation form for functions (for example +, *) and predicates (for example =, >=) may be used.

A **constant *N*-finite index term of a finite signature** is an *N*-finite index term of a finite signature which does not contain occurrences of variables.

The definition of the notion of an ***N*-finite index atomic formula of a finite signature** may be received from the definition of the notion of a *many sorted (many typed) atomic formula of a finite signature* by means of replacing the word "term of a finite signature" by the words "*N*-finite index term of a finite signature". Only *N*-finite Pascal total functions and *N*-finite Pascal total predicates of the signature, *N-reals* and *N-integers* are used in the new definition.

The definition of the notion of an ***N*-finite index first order formula of a finite signature** may be received from the definition of the notion of a *many sorted (many typed) first order formula of a finite signature* by means of replacing the word "formula" by the words "*N*-finite index formula" and the addition to the definition the possibility of the using quantifiers upon array elements.

As it is usual for the Pascal language it is allowed to insert additional pairs of brackets into N-finite index term and into N-finite index formula in a finite signature.

It is allowed type coercion: *Boolean* into *N-integer* (*true* into 0 and *false* into –1) and *N-integer* into *N-real*.

Essentially, the notion of an objective variable used in the first order logic is extended up to an array element with constant indexes. The last ones may be situated immediately after a quantifier, have free or bound occurrences in an *N*-finite index first order formula of a finite signature. That is why it is needed to complete in a natural way the definition of a **quantifier scope** as such a minimal upon the length N-finite index sub-formula of a finite signature which begins with this occurrence of a quantifier.

An occurrence of an array element *t* with constant indexes is a **bounded occurrence** if it is in a scope of a quantifier immediately after which a name of array or sub-array (with indicating constant bounds of elements in every dimension) containing the element *t* is situated. An occurrence of an array element which is not a bounded one is called a **free** one.

Let's an *N*-finite Pascal total function or an *N*-finite Pascal total predicate uses only such types as *integer* or *real* or *Boolean* and arrays of these types.

The notion of **identically true *N*-finite index formula of a finite signature** (which consists of *N*-finite Pascal total functions and *N*-finite Pascal total predicates of *N-real* or *N-integer* or *Boolean* types) may be defined as usually.

Correctness of an N-finite Pascal total function

Correctness of an *N*-finite Pascal total function body may be described by Hoare triads in the form *{A}S{B}*, where *A* and *B* are conditions, and *S* is a Pascal statement sequence [5, 7]. It may be written in the form $A \Rightarrow [B]^{\mathbf{x}}_{S1(x)}$. Here the notation $[B]^{\mathbf{x}}_{S1(x)}$ is used for the result of substitution into *B* the index terms from the list **S1**(**x**) instead of the correspondent variables from the list **x**. Both lists must have the same number of members. A name of a global variable of a statement sequence *S* is regarded as a name of an *N*-finite Pascal total function computing its value as a result of this *N*-finite Pascal total function (with the body *S*) run with the input data **x**. If an array is used as a result of a function then the name of this function coincides with the name of the array.

If all these *N*-finite Pascal total functions are included into the signature of the *N*-finite index first order logic then an assertion about full correctness of an *N*-finite Pascal statement sequence *S* may be formalized with the use of such an *N*-finite index first order logic language.

This section illustrates a new possibility of the introduced *N*-finite index first order logic of a finite signature using *N*-finite Pascal total functions and total *N*-finite Pascal predicates.

Main results

The first theorem is a practically useful generalization of an application variant of the theorem from [8]. Definitions of the **QBF** problem and P-SPACE-completeness are in [1, 2, 4]. The second theorem may be regarded as a generalization of the first one.

Theorem 1. *For every integer N (N ≥ 2) and for every finite signature of N-finite RAM or N-finite RASP total functions and total predicates containing {+, *, <} with constants from the set of all N-integers the problem of identical truth of an N-finite index first order formula (using arrays with the number of dimensions bounded by N) of such a signature is a P-SPACE-complete problem.*

Proof. An identical truth of a quantifier Boolean formula (**QBF**) which has a prenex normal form may be simulated by an *N*-finite RAM total program and sequentially by an *N*-finite RASP total program (with $N \geq 2$).

Let the logical value *true* corresponds to the number 0 and the value *false* corresponds to the number -1. In such a case logical connectives may be changed by a sequence of RAM commands:

$\neg x$	corresponds to	$-x - 1$,
$x \vee y$	corresponds to	$- x * y$,
$x \& y$	corresponds to	$(- x - 1)*(- y - 1) - 1 = (x+1)*(y+1) - 1$.

Quantifiers may be changed by loops based on *goto* statements. The length of the received *N*-finite RAM program increases relatively the initial one not more than in a polynomial under the initial QBF length times.

Hence, the formulated in the theorem problem is a P-SPACE-hard one as the **QBF** problem is a P-SPACE-complete one [1, 2, 4].

The formulated in the theorem problem belongs to the class **P-SPACE** because all *N-finite* RAM or *N-finite* RASP total program commands belong to **FP**. Besides, during such a program run the memory may be bounded by a polynomial under the input data length because the looping does not need the use of an additional array element.

Theorem 2. *For every integer N ($N \geq 2$) and for every finite signature of N-finite non-recursive Pascal total functions and N-finite non-recursive Pascal total predicates containing {+, *, <} with constants from the set of all N-reals and all N-integers the identically true problem of an N-finite index first order formula (using arrays with the number of dimensions bounded by N) of such a signature is a P-SPACE-complete problem.*

Proof. The number of *non-recursive* *N*-finite Pascal total functions and *non-recursive* *N*-finite Pascal total predicates are finite ones. Every *N*-finite Pascal total function and every *N*-finite Pascal total predicate belongs to **FP** because it has a finite domain and its values may be computed as a preliminary data base.

An *N*-finite Pascal total function is a generalization of an *N*-finite RAM total program. That's why the formulated in the theorem problem is a P-SPACE-hard one. This problem belongs to the class **P-SPACE** because an *N*-finite Pascal total program may be interpreted by an *N*-finite RAM program. Of course, the description of such a translator is very large and includes the full description of the Pascal language as a part.

Note. The predicate "<" allows easily to define the relation "=". But the author does not know a short definition of the predicate "<" in the terms of "=". That is why the relation "<" is used in the conditions of the above theorems.

The set of all *N*-finite index first order formulas of the signature from the theorem 2 may be named an *N*-finite arithmetic of such a signature.

Conclusion

Models of a mathematical notion of a program processing only finite short data types are introduced for adequate description of a computer program. These models process only data of the computer types *real, integer and Boolean* with the bounded notation length.

Mathematical notions of an *N-finite* RAM program and an *N-finite* RASP program are introduced as generalization of a RAM program and a RASP program for computation of functions and predicates of a finite arithmetic signature. The notion of an *N-finite* Pascal program is introduced as an extension of an *N-finite* RAM program up.

An extension of a first order predicate formula is done on the base of *N-finite* data. Such an extension allows describing mathematical properties of total Pascal function or predicate. P-SPACE-completeness of the truth checking for these formulas with non-recursive functions and predicates is proved.

The introduced formulas are useful for formulation of a program correctness condition by means of Hoare triades.

Acknowledgement

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).

Bibliography

1. Aho A.V., Hopcroft J.E., Ullman J.D. The design and analysis of computer algorithms. (Addison-Wesley Publishing Company Reading, Massfchusetts, 1976.
2. Du D.Z., Ko K.I. Theory of Computational Complexity. A Wiley-Interscience Publication. John Wiley & Sons, Inc. 2000.
3. Forsyth R.S. Pascal at Work and Play. An introduction to computer programming in Pascal. Chapman & Hall. London, New-York. 1982.
4. M.R. Garey and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, New York, 1979.
5. Hoare C.A.R. A note on the for statement. BIT (Sver), 1972, vol.12, No 3.
6. Kleene S.C. Mathematical Logic. A Wiley-Interscience Publication. John Wiley & Sons, Inc., 1967.
7. Kosovskii N.K. Elements of mathematical logic and its applications to the theory of sub-recursive algorithms. Leningrad University Press, Leningrad, 1981. (In Russian)
8. Kosovskiy N.K. Pspace-completeness of finite order predicate logic over finite domain. 3 International Conference "Smirnov's lectures". Moscow, 2001, p. 44. (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 Computational Complexity of Algorithms.

WORLDDYN AS THE TOOL FOR STUDY OF WORLD DYNAMICS WITH FORRESTER’S MODEL: THEORY, ALGORITHMS, EXPERIMENTS

Olga Proncheva

Abstract: *Forrester´s model (FM) is a system of nonlineal differential equations constructed at far 70-s. Since that time FM remains a classical example of applications of system dynamics principles. This paper shortly describes the theory, algorithms and program WorldDyn for end-users being interested to meet themselves with FM. Moreover the program allows to model undefiniteness presented in the form of additive or multiplicative noise. By the moment there is no any free-share software having these possibilities.*

Keywords: *Forrester model, numerical analysis, noise-immunity, applied software.*

1. Classic Forrester’s model

1.1. Principles of system dynamics

System dynamics is based on two main principles. First of all, equations of the same type are written for all variables:

$$\frac{\mathrm{dy}}{\mathrm{dt}} = \mathrm{y}^{+} - \mathrm{y}^{-}. \quad (1)$$

Here y+ is the positive rate of variable change (it includes all factors calling the rise of the variable y), y– is the negative rate of variable change (it includes all factors calling decrease of the variable y).

Thereafter it is supposed that all rates (the positive and negative ones) could be presented in the form of function compositions, which depend on one factor (combination of main variables):

$$y^{\pm} = g(y_1, y_2, \dots, y_n) = f(F_1, F_2, \dots, F_k) = f_1(F_1) f_2(F_2) \dots f_k(F_k). \quad (2)$$

1.2. Forrester’s model

Roman Club is nongovernmental organization, which joins political und scientific personalities and is working on modeling of World Crisis. At far 1970 the elite Roman club asked prof. J. Forrester from MIT to develop a model of world dynamics. Speaking world dynamics we mean the dynamic interactivity of the main macroeconomics variables. The first version of the model named “World-1” was presented in 4 weeks and next year the corrected version “World-2” [2] was accepted as the classical J. Forrester model. In spite of its long history the J. Forrester model retains its actuality being the basis for modern models, so we can say, that this model is actual even today, in spite of elderly age. Such models, based on J. Forrester’s model, can predict crisis and sometimes avoid it. So, such models are very important.

As far classical model, J. Forrester in his work [2] saw five main problems, because of which the World Crisis can appear. It is overpopulation of our planet, lack of basis resources, the critical level of pollution, food shortages and industrialization and the related industrial growth. He tied a single variable with each of these issues. So, we have a five-level system, on which is built the structure of the system:

- Population (P);
- Pollution (Z);
- Natural resources (R);
- Fixed capital (K);
- Capital investment in agriculture fraction (X).

For system level J. Forrester brought the following differential equation:

$$\frac{dP}{dt} = P(c_B B_C B_P B_F B_Z - c_D D_C D_P D_F D_Z) \quad (3)$$

$$\frac{dK}{dt} = c_K P K_C - \frac{K}{T_K} \quad (4)$$

$$\frac{dX}{dt} = \frac{X_F X_Q - X}{T_X} \quad (5)$$

$$\frac{dZ}{dt} = P Z_K - \frac{Z}{T_Z} \quad (6)$$

$$\frac{dR}{dt} = -P R_C \quad (7)$$

Here he used tabulated functions (with linear interpolation) $(B_C, B_P, B_F, B_Z, D_C, D_P, D_F, D_Z, K_C, X_F, X_Q, Z_K, R_C)$ and constants:

c_B=0,04 (normal fertility rate), c_D=0,028 (normal death rate), c_K=0,05 (normal rate of capital), T_K=40 (time of depreciation main funds), T_X=15 (time of depreciation agricultural funds), t_N=1970 (initial year), P_N=3,6·10^9 (population in initial year), X_N=0,3 (capital investment ratio in agriculture in initial year), Z_N=3,6·10^9 (pollution in initial year).

Initial data are:

$$t_0 = 1900, P_0 = 1{,}65 \cdot 10^9, K_0 = 0{,}4 \cdot 10^9, X_0 = 0{,}2, Z_0 = 0{,}2 \cdot 10^9, R_0 = 900 \cdot 10^9.$$

Standard pollution Z_N is numerically equal to the population, and R_0 was taken on the assumption that resource at a constant rate of consumption (equal to the rate of consumption in 1970) should be sufficient for 250 years.

In addition Forrester imputed such variables, as consumption (F) and material level of living (C):

$$F = F_X F_P F_Z \quad (8)$$

$$C = K_p \frac{1 - X}{1 - X_0} E_R \quad (9)$$

Here F_X, F_P, F_Z, K_p, E_R are tabulated functions.

1.3. Experiment 1: instability

If we try to solve equations (3)-(7), we will get following results. The behavior of the model parameters is shown in Figure1:

One can see (Figure 1), that after a period of growth, the population P begins to decline since 2020. Non-renewable natural resources in 2100 are less than ⅓ of the original stock. Pollution reaches its maximum in 2050, about 6 (more precisely, 5.8) times exceeding the standard level, then drops due to the general decline of industry and population decline. Material standard of living reaches its maximum about 2000, and then decreases.

The reason of it is resource depletion. Reducing the supply of resources R causes lower material standard of living C. This causes increase mortality and reduce investment. And, finally, we have a sharp population decline and fall of industrial production (of funds). J. Forrester tried to change the original settings in order to avoid the crisis, but every time the crisis arose.

1.4. Experiment 2: Global equilibrium

So Forrester suggested changing some constants in the model that means political reforms in 1970 (the year, when he built his model). More specially, he offered [2]:

1. To decrease natural resources usage rate in 4 times (as compared with 1970 year);
2. To decrease pollution generation in 2 times;
3. To decrease capital investment on 40%;
4. To decrease birth rate on 30%.

In the Figure 2 we can see the behavior of main variables under these assumptions:

In this case we come to so-called “global equilibrium”. Behavior of model is improved, but we don’t solve the problem. Nevertheless we can adapt to situation.

In this paper we will discuss “stable” model.

1.5. Forrester’s followers

Forrester’s model was developed more than 40 years ago, so many researches worked with this model.

D. Meadows [5] developed World-3. He suggested including more than one variable in each “problem” sector. But unfortunately he had too little data for his model, so he tried to retrieve data according to Forrester’s model.

In USSR and Russia there were Forrester’s followers too. Matrosov [4] suggested including new factors, such as biomass of the Earths, scientific-and-technological advance, political tension. He proved, that there are stationary solutions and they are stable. Egorov [1] influenced on material standard of living, pollution ratio, food coefficient. And he supposed that there is technology of utilization and recovery resources, artificial cleaning pollution and investment in agriculture can be changed. S. Makhov [3] removed from the model such variables, as pollution and capital investment in agriculture fraction, but includes energy resources and education level.

By the moment there is no an accessible end-user program based on this model and our goal was the development of such a program with a comfortable graphical interface. And no one has investigated an influence of white noise on the model. We will try to do this analysis.

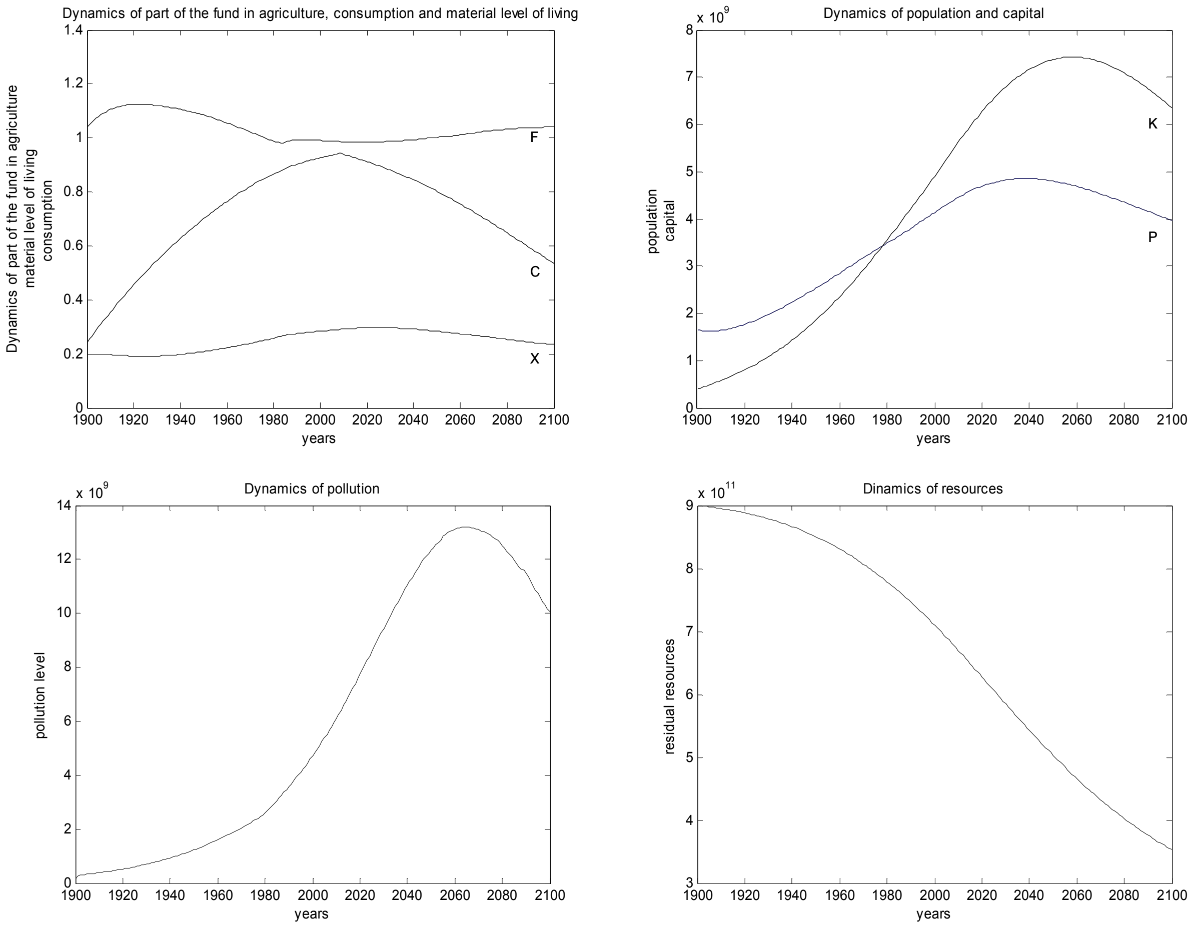


Figure 1. Behavior of main variables in Forrester’s model

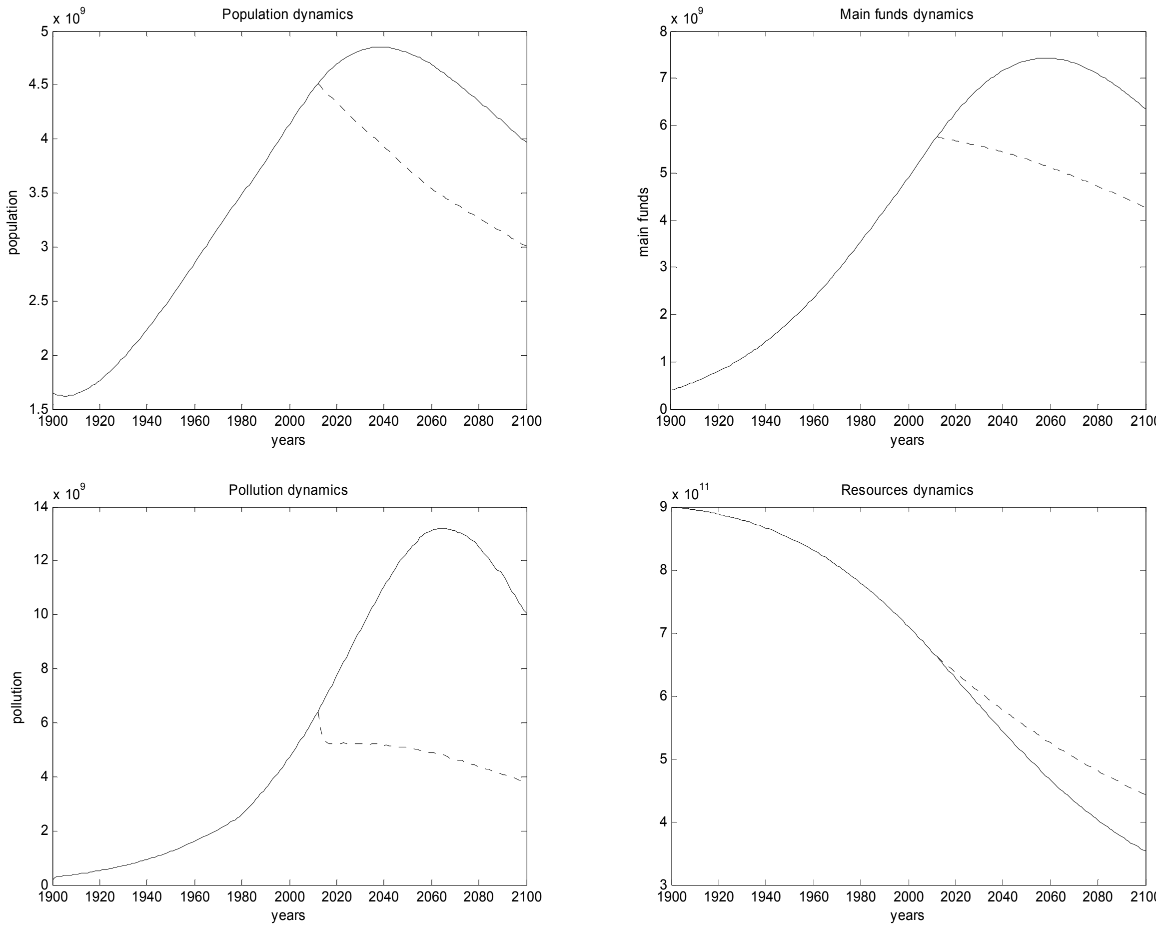


Figure 2. Behavior of main variables under assumptions 1-4. Dotted line: results of modeling under assumptions, firm line: initial behavior

2. Forrester’s model with noise

2.1. What means noise?

In this research we affect original Forrester’s model with additive and multiplicative noise. Under additive noise we understand external influences on the system (outside the scope of this level of consideration) that cannot be controlled. Under additive noise we understand disturbance within the system. We expect that influence of multiplicative noise will be less than influence of additive noise.

2.1. Modeling with additive noise

We appended white additive noise in 1970, in other worlds, on forecast. In this case the model before 1970 is setted by equations (3-7).

After 1970 the model is:

$$\frac{dP}{dt} = P(c_B B_C B_P B_F B_Z - c_D D_C D_P D_F D_Z) + \vartheta \tag{10}$$

$$\frac{dK}{dt} = c_K P K_C - \frac{K}{T_K} + \theta \tag{11}$$

$$\frac{dX}{dt} = \frac{X_F X_Q - X}{T_X} + \mu \tag{12}$$

$$\frac{dZ}{dt} = P Z_K - \frac{Z}{T_Z} + \sigma \tag{13}$$

$$\frac{dR}{dt} = -PR_C + \tau \tag{14}$$

where ϑ, θ, μ, σ, τ – stationary white noise.

When we find out the most powerful and the most sensitive variable, we append white noise only in one equation.

The noise was constructed as follows.

Signal strength was taken during the teaching phase, that is, for each variable was considered integral in the form:

$$\frac{1}{71}\int_{1900}^{1970} f^2(t)dt$$

Next, white noise was generated; in each time this value was multiplied by the noise signal power for each of the variables.

After that the system was solved 100 times, we got 100 "noise" functions. For each time point we calculated the arithmetic mean, the resulting function is the mathematical expectation of the process.

As a measure of stability, we used the relative standard deviation:

$$\sigma_i^{rat} = \frac{1}{131} * \sum_{t=1970}^{2100} \frac{1}{f_i(t)} * \sqrt{\sum_{k=1}^{100} (f_i^k(t) - f_i^{mean}(t))^2 * \frac{1}{100}} \tag{15}$$

where

$f_i(t)$ - noiseless value of the i-th variable (i = $\overline{1,5}$) - the number of variable in succession);

$f_i^k(t)$ – value at time t of i-th variable in the k-th realization;

$f_i^{mean}(t)$ – the mathematical expectation of k-th variable as defined above.

First of all we studied noise influence on the system, when noise affects all variables. We considered the case with 20% noise level. This level is close to the critical value, when some solutions become diverge. Figure 3 presents the results of modeling for all macroeconomics variables. There are 3 lines on the figure: thin uninterrupted line is the initial function, thick line is the forecast, and thin dotted line is the worst function.

One can see that the forecast is very close to the initial line, which reflects unnoised function. It means that our model is stable to the 20% noise. In this case all functions converge and the relative root-mean-square deviation for every variable is equal:

- 3,52% for population;
- 4,26% for main funds;
- 4,38% for agriculture;
- 6,00% for pollution;
- 6,45% for resources.

Generally speaking, the Forrester model is very stability to the noise, which acts on the forecast period. Even when the noise level reaches 50% we have 69% convergent functions [6]. This result also can be considered as the very good one.

We also have done some experiments, when noise affects only one of variables. This way we wanted to find out the most sensitive and the most influential variable. When we talk about the most sensitive variable, we mean variable, which reacts most strongly to a noisy of the other variables. When we talk about the most influential variable, we mean variable, noisy of which influence the most other variables. We found out, that the most sensitive variable is pollution and the most influential variable is resources. Figure 4 illustrates the population dynamics given the influence of different variables.

In the case, when noise affects only resources, the relative root-mean-square deviation for every variable is equal:

- 0,47% for population;
- 0,78% for main funds;
- 0,41% for agriculture;
- 1,95% for pollution;
- 0,71% for resources;

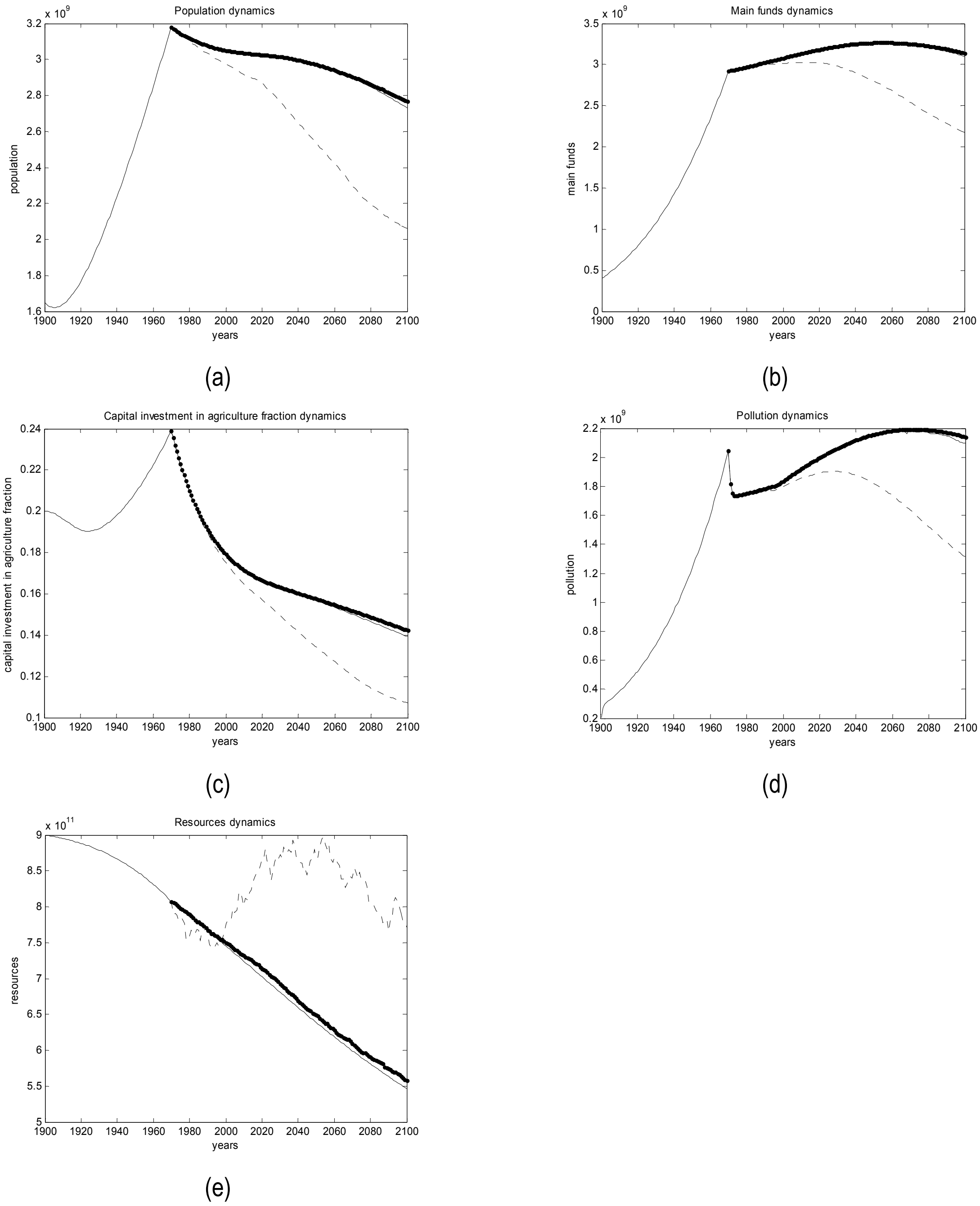


Figure 3. Results of modeling (a) population dynamics; (b) main funds dynamics; (c) dynamics of capital investment in agriculture fraction (d) pollution dynamics; (e) resources dynamics.

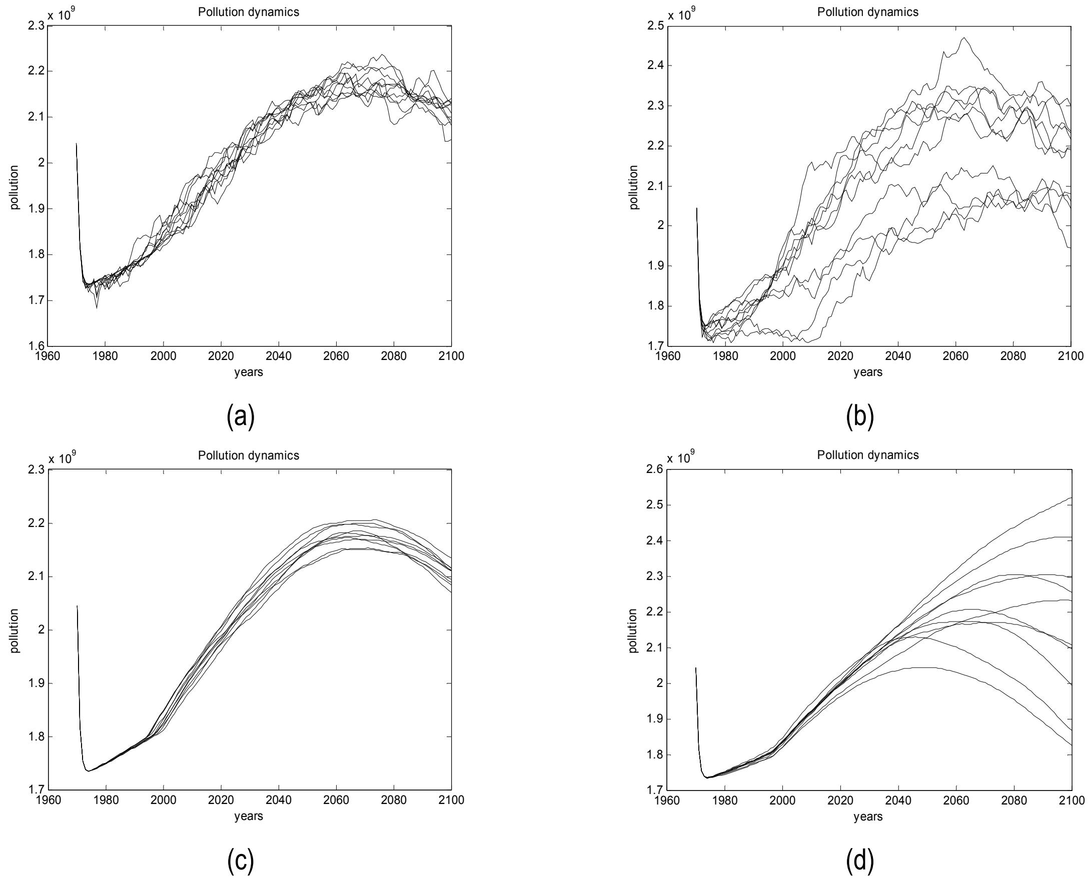


Figure 4. Results of modeling pollution dynamics (a) noise affects only population; (b) noise affects only main funds; (c) noise affects only pollution; (d) noise affects only resources

2.3. Modeling with multiplicative noise

We appended white multiplicative noise in 1970, in other worlds, on forecast. In this case the model before 1970 is setted by equations (3-7), after 1970 the model is:

$$\frac{dP}{dt} = P(c_B B_C B_P B_F B_Z - c_D D_C D_P D_F D_Z) * (1 + \vartheta) \tag{16}$$

$$\frac{dK}{dt} = (c_K P K_C - \frac{K}{T_K}) * (1 + \theta) \tag{17}$$

$$\frac{dX}{dt} = \frac{X_F X_Q - X}{T_X} * (1 + \mu) \tag{18}$$

$$\frac{dZ}{dt} = \left(P Z_K - \frac{Z}{T_Z}\right) * (1 + \sigma) \tag{19}$$

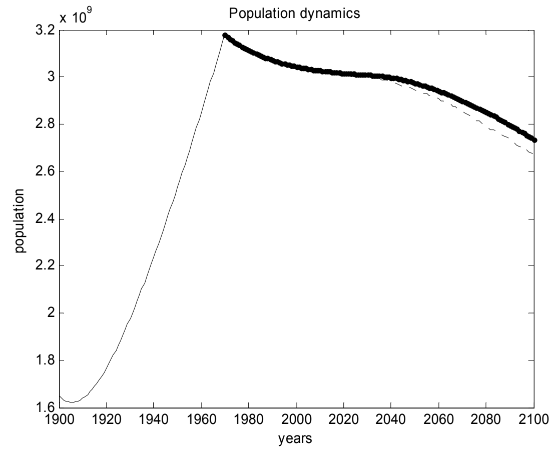
$$\frac{dR}{dt} = -P R_C * (1 + \tau) \tag{20}$$

where ϑ, θ, μ, σ, y,τ – stationary white noise. Each of these noises has power, it means, that we multiply it by its power.

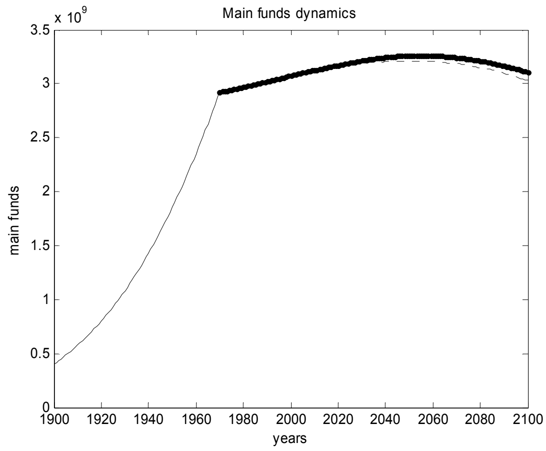
After that the system was solved 100 times, we got 100 "noise" functions. For each time point we calculated the arithmetic mean, the resulting function is the mathematical expectation of the process.

As a measure of stability, we used (15).

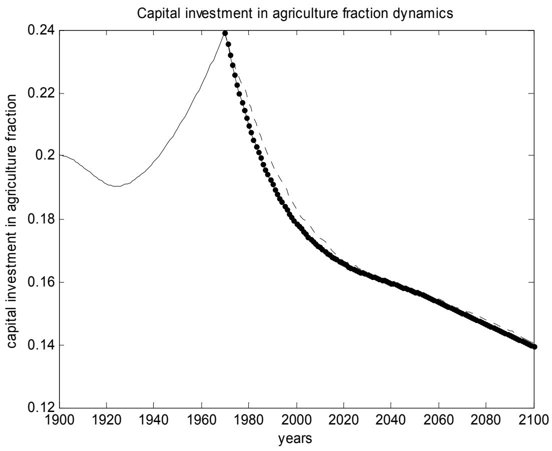
We studied noise influence on the system, when noise affects all variables. We considered the case with 50% noise level. This level is close to the critical value, when some solutions become diverge. Figure 5 presents the results of modeling for all macroeconomics variables. There are 3 lines on the figure: thin uninterrupted line is the initial function, thick line is the forecast, and thin dotted line is the worst function.



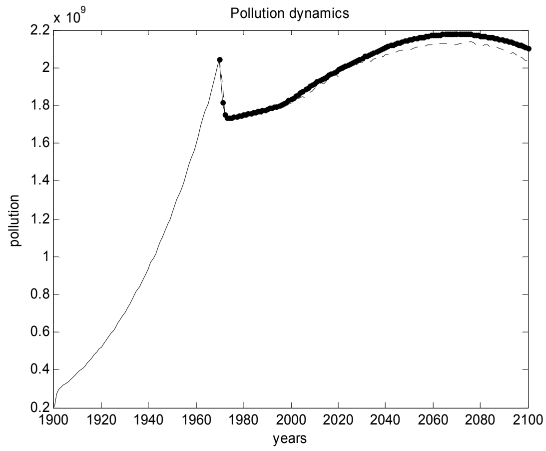
(a)



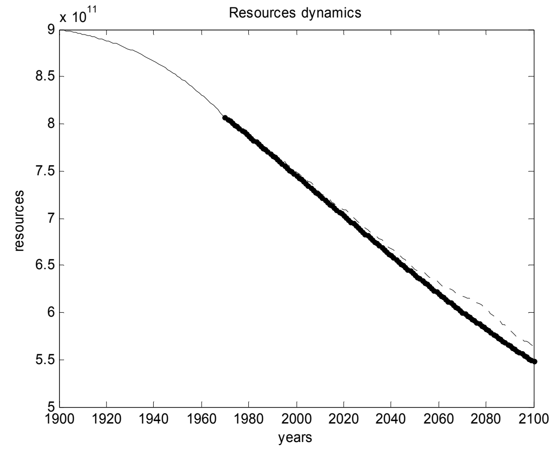
(b)



(c)



(d)



(e)

Figure 5. Results of modeling (a) population dynamics; (b) main funds dynamics; (c) dynamics of capital investment in agriculture fraction (d) pollution dynamics; (e) resources dynamics

We can see that the model is more stable to multiplicative noise, than to additive noise. In this case all functions converge and the relative root-mean-square deviation for every variable is equal:

- 0,23% for population;
- 0,28% for main funds;
- 0,48% for agriculture;
- -,50% for pollution;
- 0,49% for resources.

In the case with multiplicative noise we find out (as in the case with additive noise), that the most sensitive variable is pollution, and the most influential variable is resources.

Figure 6 illustrates the population dynamics given the influence of different variables.

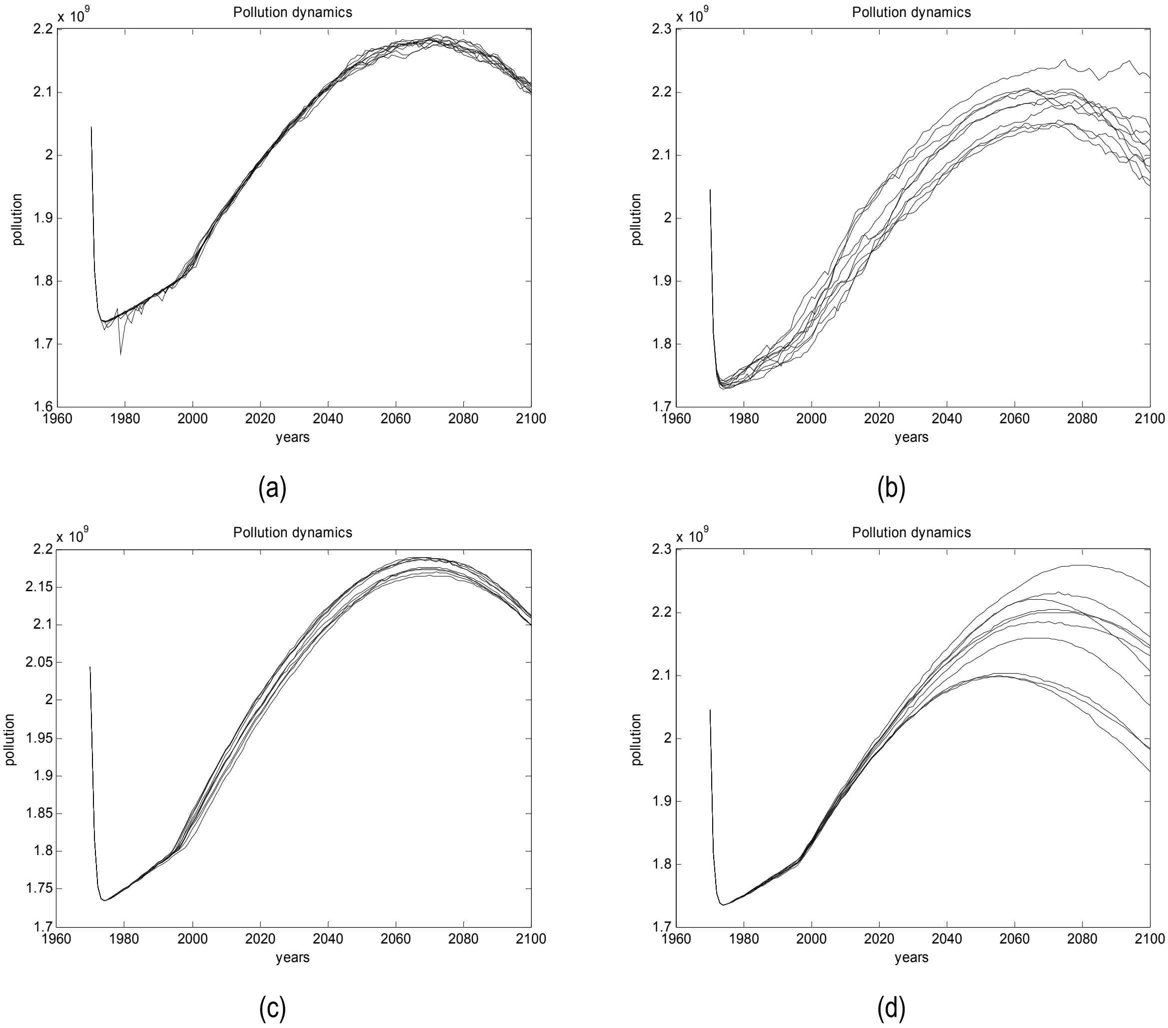


Figure 6. Results of modeling pollution dynamics (a) noise affects only population; (b) noise affects only main funds; (c) noise affects only pollution; (d) noise affects only resources

In the case, when noise affects only resources, the relative root-mean-square deviation for every variable is equal:

- 0,91% for population;
- 0,93% for main funds;
- 1,03% for agriculture;
- 1,13% for pollution;
- 1,83% for resources.

2.4. Comparative analysis

If we compare cases with additive and multiplicative noise, we can draw a conclusion, that, as we expected, multiplicative noise affect the system much less, than additive nose. And in both cases we get identical results about the most and the less sensitive and influential variable. We found out, that the most influential variable is resources. It means that we should pay large attention on resources, if we want to save stability of target system.

3. Research of crisis situation

3.1. What means "Research of crisis situation"?

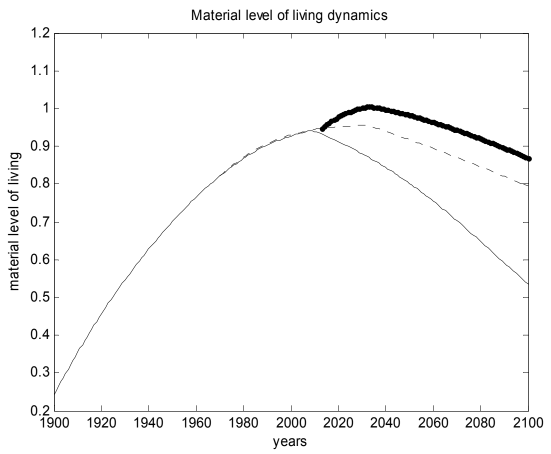
By the instrumentality of Forrester's model, it is possible to predict world crises. And we can change parameters, as well as Forrester changed, to avoid crises. Parameter's modifications mean different reforms, which can influence parameters of the model, such as birth rate or pollution rate restrictions.

3.2. Experiment 1

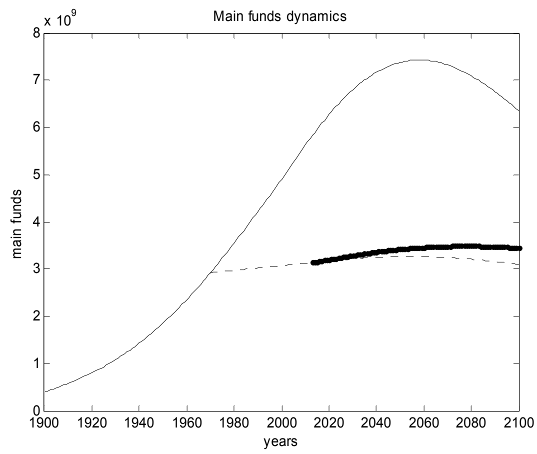
For example, in addition to Forrester changes in 1970 we can append some changes in 2013 year. We suggest the following (in compare with 1970 year):

- To decrease birth rate on 10%;
- To decrease pollution on 20%.

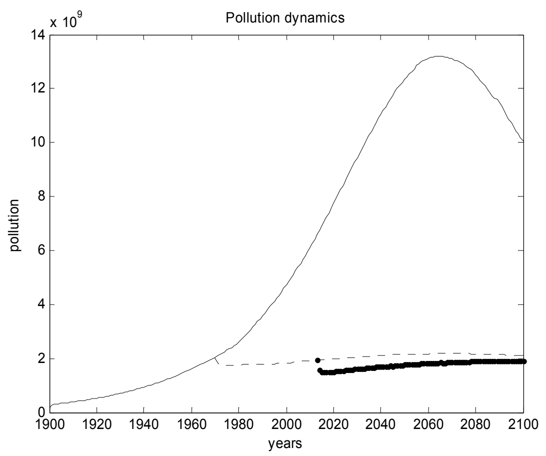
In this case we get (figure 7):



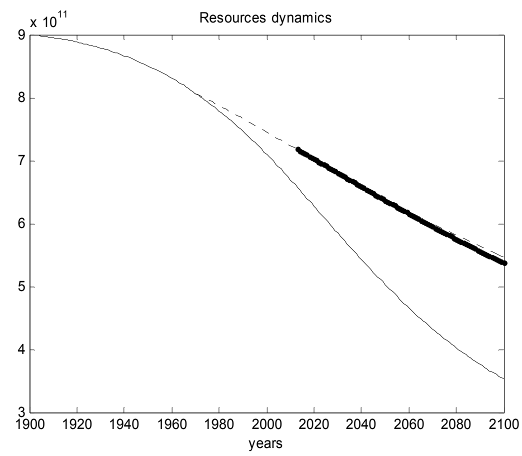
(a)



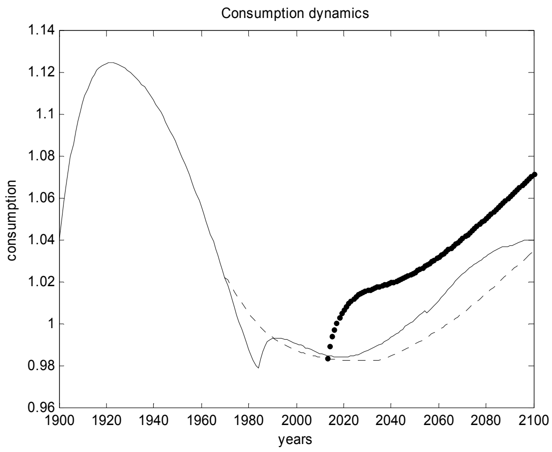
(b)



(c)



(d)



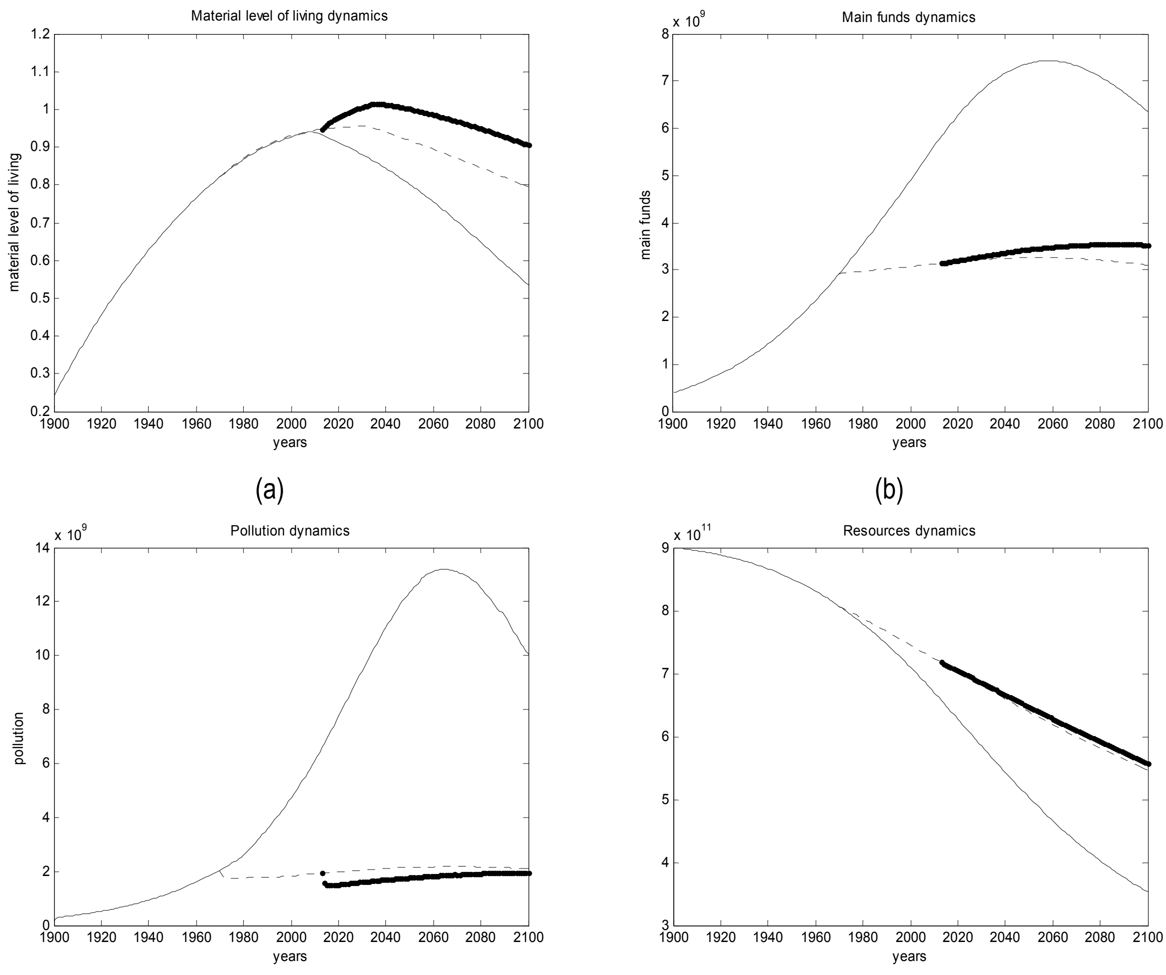
(e)

Figure 7. Results of modeling (a) material level of living dynamics; (b) main funds dynamics; (c) pollution dynamics; (d) resources dynamics; (e) consumption dynamics. Thin uninterrupted line is the initial function, thick line is the function after first break, and thin dotted line is the function after second break

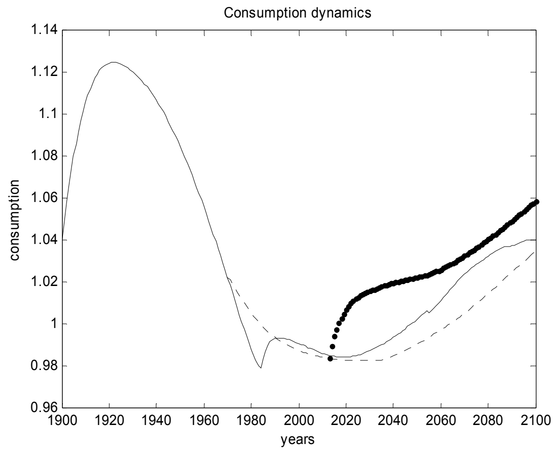
It is easy to see, that we improve material level of living and consumption, but don't change situation with resources. So we should continue experiments with parameters.

3.3. Experiment 2

In addition to previous reforms we suggest to decrease resources consumption on 20%. The results one can see in figure 8:



(c) (d)



(e)

Figure 8. Results of modeling (a) material level of living dynamics; (b) main funds dynamics; (c) pollution dynamics; (d) resources dynamics; (e) consumption dynamics. Thin uninterrupted line is the initial function, thick line is the function after first break, and thin dotted line is the function after second break

Unfortunately situation with resources doesn't change. But consumption is much better, even if compare it with initial dynamics.

4. WORLDDYN program

4.1. The development environment and the structure of the program

This program was developed in MATLAB [8]. The program is user-friendly, even if user doesn't have any skills in programming, he can work with WORLDDYN [7]. There are a several forms for each kind of experiments: adding additive or multiplicative noise, change parameters. One can see results of modeling in numerical forms (dispersions and percent of convergent functions) and in graphic form (figures).

4.2. Interface

A user can complete the experiments with noise. To make such an experiment in initial data the user should open "Options"-> "Noise settings (initial data)" (figure 9).

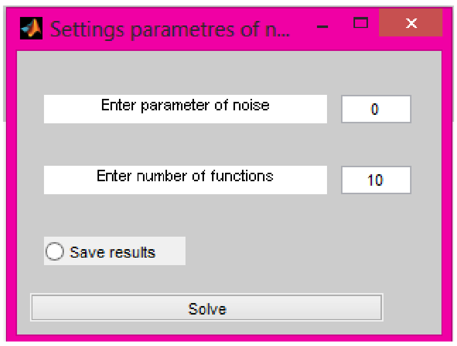


Figure 9. Menu for noise definition in the initial data

After that he chooses the noise level (parameter of noise) and the number of noise functions. To save the results the user points "Save results". The forecast will be save as "change_initial_*(name of variable)_expectation_*(noise level)noise_*(the number of functions)functions_*(the percent of converge functions)%_converge". And finally the user enters "Solve".

In order to complete the experiment with noise related with all variables on a forecast stage the user should open "Options"-> "Noise settings (forecast)" (figure 10):

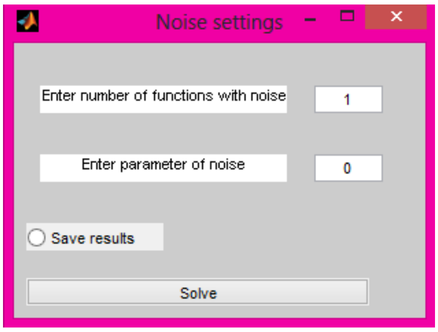


Figure 10. Menu for noise definition on the forecast stage

After that he chooses the noise level (parameter of noise) and the number of noise functions. To save the results the user points "Save results". The forecast will be saved as "change_all_*(name of variable)_expectation_*(noise level)noise_*(the number of functions)functions_*(the percent of converge functions)%_converge". And finally, the user enters "Solve".

In order to complete the experiment with noise related with one given variable on a forecast stage the user should open "Options"-> "Noise settings (only one function)" (figure 11):

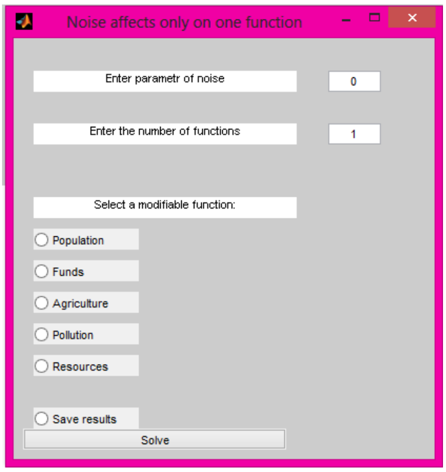


Figure 11. Menu for noise definition on the forecast stage (one variable)

After that he chooses the noise level (parameter of noise), the number of noise functions, and the name of modifiable variable. To save the results the user points "Save results". The expectation will be save as "change_only_*(name of modifiable variable) _*(name of variable)_expectation_*(noise level)noise_*(the number of functions)functions_*(the percent of converge functions)%_converge". And, finally, he can enter "Solve".

When the program World-Dyn works it can inform the user:

- About model solution (figure 12);
- About final of calculation (figure 13);
- About errors (figure 14 a, b).

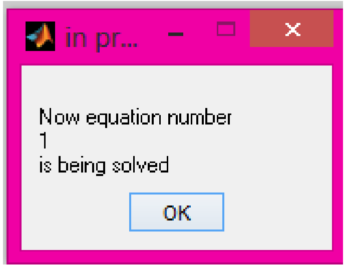


Figure 12. About model solution

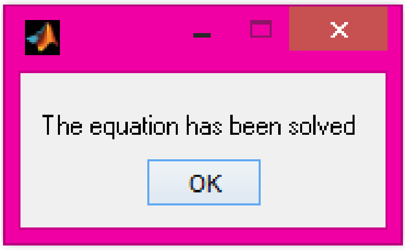


Figure 13. About final of calculation

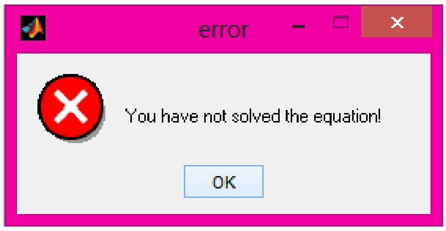


Figure 14. a) About error

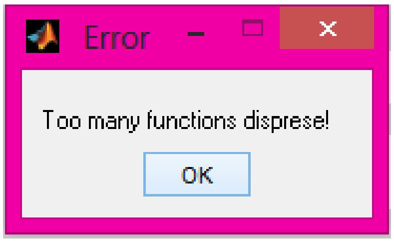


Figure 14. b) About error

4.3. Numerical methods

We used two algorithms in our research: Runge-Kutt method for modeling noise system and Adams-Bashwort-Multon method for modeling original system and system with modified parameters without noise. Consider each of these methods closer.

Runge-Kutt method

The Runge-Kutta method is as follows.

The following values are given for ordinary differential equations:

$$k_1 = f(t_n, u_n);$$

$$k_2 = f(t_n + \alpha_2\tau, u_n + \tau\beta_{21}k_1)$$

$$\dots$$

$$k_r = f(t_n + \alpha_r\tau,\ u_n + \tau(\beta_{r1}k_1 + \dots + \beta_{r,r-1}k_{r-1}))$$

$$u_{n+1} = u_n + \tau(\gamma_1 k_1 + \dots + \gamma_r k_r)$$

The factors determining the specific method can be presented as a Butcher table (table 1):

Table 1

0						
α_2	β_{21}					
α_3	β_{31}	β_{32}				
…	…	…	…			
α_r	β_{r1}	β_{r2}	…	$\beta_{r,r-1}$		
	γ_1	γ_2	…	γ_{r-1}	γ_r	

There is implemented method of third order accuracy in the program.

Adams method

Suppose we know the approximate solution of some of the computational grid nodes: $t_n, \dots, t_{n-m}$.. In the neighborhood of these nodes we replace the function of the interpolation polynomial, written in the form of Newton's:

$$f(t) = f(t_n) + f(t_n, t_{n-1})(t - t_n) + f(t_n, t_{n-1}, t_{n-2})(t - t_n)(t - t_{n-1}) + \cdots$$

In order to compute the solution at n +1, we write it in integral form:

$$u_{n+1} = u_n + \int_{t_n}^{t_{n+1}} f(t, u(t))dt = \int_{t_n}^{t_{n+1}} f(t)dt$$

4.4. Modeling protocol

User can see results in two forms: numerical and graphic. If user chooses graphic form, he can choose two modes: see all noise functions (figure 15, a) or 3 functions: the worth (the highest dispersion), mathematic expectation and unnoise function (figure 15, b). If user chooses numerical form, he will see next forms (figure 16 a, b):

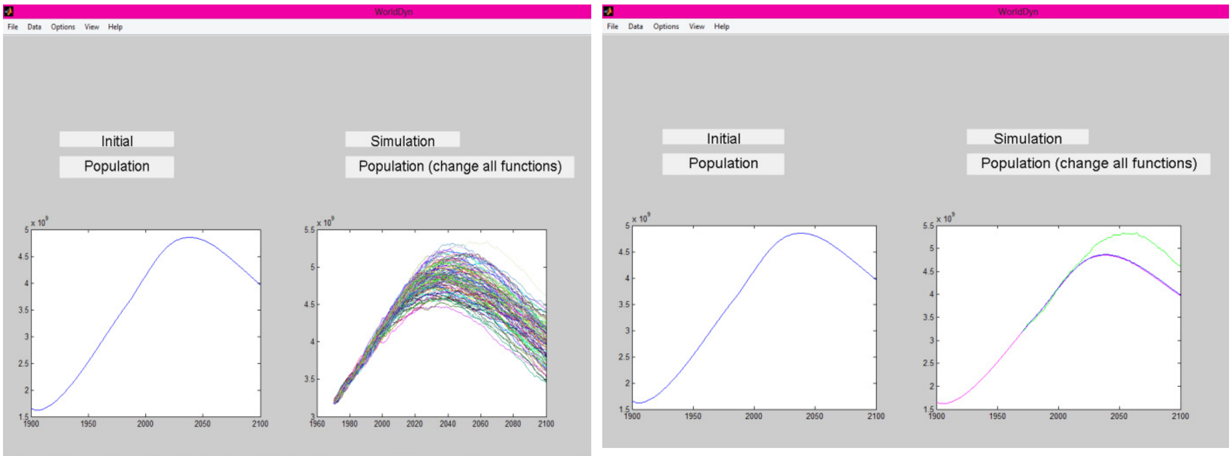


Figure 15. Graphic mode. a) All noise functions. b) 3 functions: the worth (the highest dispersion), mathematic expectation and unnoise function

Figure 16. Numerical mode a)Dispersion, b) Percent of converge functions

4.5. Rules of statement

User should choose the purpose of his research. Hence he will choose regimes of modeling and output.

If user has some problems with WORLDDYN, he can use help. There is workbook in Russian, soon will workbooks in English and Spanish.

5. Conclusion

In the paper we have studied the stability of the classical Forrester model to additive and multiplicative noise. The experiments show that

- Additive noise causes the essentially stronger effect then the same multiplicative noise on the stage of forecast;
- The most influential variable is resources; its changes provoke the strongest reaction of the model. The most sensitive variable is pollution.

Bibliography

1. Egorov, V., Kallistov N., Mitrofanov, N., Piontkovsky, V. Mathematical model of global development: a critical analysis of models of nature. - Gidrometeoizdat, 1980. - 192. (rus).
2. Forrester J. World Dynamics. - Moscow: AST, 2003. - 379 p. (rus)
3. Makhov S. Long-term trends and projections in terms of a new model of world dynamics/forecast and modeling the dynamics of the global crisis/Ed. A. Akayev, A. Korotayev, G. Malinetskii/Future Russia. M.: LCI, 2010, P. 262–276 (rus)
4. Matrosov V. Matrosov I. Global modeling taking into account the dynamics of biomass and scenarios for sustainable development / new paradigm of development in Russia (Comprehensive studies on sustainable development). - Moscow: Academia, MGUKI, 1999. - S. 18 – 24 (rus).
5. Petrov I., Lobanov A. "Lectures in Computational Mathematics." - M: The Internet University of Information Technology; BINOM. Laboratory of Knowledge, 2006. - 523 p. (rus)
6. Proncheva O., Alexandrov M., Makhov S. Testing stability of classical Forrester model to initial data and additive noise / in. Artificial Intelligence Driven Solutions to Business and Engineering Problems / Eds. Galina Seltak, Mikhail Alexandrov, Krassimir Markov. - Rzeszow-Sofia: ITHEA, 2012, V.26, P. 201 - 209.
7. Proncheva O. Program WORLDDYN based on Forrester model of world dynamics / in. Artificial Intelligence Driven Solutions to Business and Engineering Problems / Eds. Galina Seltak, Mikhail Alexandrov, Krassimir Markov. - Rzeszow-Sofia: ITHEA, 2012, V.27, P. 16 - 20.
8. Vasilev A. "Matlab. Tutorial. A Practical Approach". - Moscow: Science and Technology, 2012. - 448 p. (rus).

Authors' Information

***Olga Proncheva** – Bachelor, Russian Presidential Academy of national economy and public administration; Prosp. Vernadskogo 82, bld. 1, Moscow, 119571, Russia; Moscow Institute of Physics and Technology (State University); Institutskii per 9., Dolgoprudny, Moscow Region, 141700, Russia; e-mail: olga.proncheva@gmail.com*

Major Fields of Scientific Research: mathematical modeling, world economy

COMPARISON OF DIFFERENT WAVELET BASES IN THE CASE OF WAVELETS EXPANSIONS OF RANDOM PROCESSES

Olga Polosmak

Abstract: *In the paper wavelets expansions of random processes are studied. The matter is that although it is enough information for wavelets expansions of deterministic functions, for random processes such theory is weak and it should be developed. The paper investigates uniform convergence of wavelet expansions of Gaussian random processes. The convergence is obtained under simple general conditions on processes and wavelets which can be easily verified. Applications of the developed technique are shown for several wavelet bases. So, conditions of uniform convergence for Battle-Lemarie wavelets and Meyer wavelets expansions of Gaussian random processes are presented. Another useful in various computational applications thing is the rate of convergence, especially if we are interested in the optimality of the stochastic approximation or the simulations. An explicit estimate of the rate of uniform convergence for Battle-Lemarie wavelets and Meyer wavelets expansions of Gaussian random processes is obtained and compared.*

Keywords: *random processes, wavelets expansion, uniform convergence, Battle-Lemarie wavelets, Meyer wavelets, Gaussian processes.*

ACM Classification Keywords: *G.3 Probability and Statistics - Stochastic processes*

Introduction

Wavelet analysis is an exciting effective method for solving difficult problems in mathematics, physics, economics, medicine and engineering.

The most actual issues of application of wavelet analysis related with signal processing and simulation, audio and image compression, noise removal, the identification of short-term and global patterns, spectral analysis of the signal. From a practical point of view, multiresolution analysis provides an efficient basis for the expansion of stochastic processes. Wavelet representations could be used to convert the problem of analyzing a continuous-time random process to that of analyzing a random sequence, which is much simpler. This approach is widely used in statistics to estimate a curve given observations of the curve plus some noise, in time series analysis for smoothing functional data, in simulation studies of various functionals defined on realizations of a random process, etc.

Recently, a considerable attention was given to wavelet orthonormal series representations of stochastic processes. Some results, applications, and references on convergence of wavelet expansions of random processes in various spaces can be found in [Atto et al., 2010; Bardet et al., 2010; Didier et al., 2008; Kozachenko et al., 2011, 2013; Kozachenko, Polosmak, 2008], just to mention a few.

In the paper we study uniform convergence of wavelet decompositions which is required for various practical applications (but most known results in the open literature concern the mean-square convergence of wavelets expansions). So we consider stationary Gaussian random processes $\mathbf{X}(t)$ and their approximations by sums of wavelet functions

$$\mathbf{X}_{n,\mathbf{k}_n}(t) := \sum_{|k|\le k_{0'}} \xi_{0k}\phi_{0k}(t) + \sum_{j=0}^{n-1}\sum_{|k|\le k_j} \eta_{jk}\psi_{jk}(t), \quad (1)$$

where $\mathbf{k}_n := (k_{0'}, k_0, ..., k_{n-1})$, functions $\phi_{0k}(t), \psi_{jk}(t)$ are wavelet bases (in the papper we consider Battle-Lemarie and Meyer wavelets).

In direct numerical implementations we always consider truncated series like (1), where the number of terms in the sums is finite by application reasons (this makes it possible to find an explicit estimate of the rate of uniform convergence for wavelets expansions of random processes).

The rate of convergence is very useful notion in various computational applications. But this question has been studied very little.

In this paper our focus is on the Battle-Lemarie and Meyer wavelet bases. This is done with the aim to show that all our results are not only theoretical, but they can be used in practice. Using the program Wolfram Mathematica, we get the convergence rate for Battle-Lemarie and Meyer wavelet decompositions of Gaussian random processes.

The organization of this article is the following. In the second section we introduce the necessary background from wavelet theory and a theorem on uniform convergence in probability of the wavelet expansions of stationary Gaussian random processes, obtained in [Kozachenko et al., 2011]. In the third section we give some notions about the Meyer wavelet bases and obtain conditions of uniform convergence for this wavelets. The next section contains the rate of convergence in the space $C([0,T])$ of Meyer wavelet decompositions of stationary Gaussian random processes. In the section 5 we give some notions about the Battle-Lemarie wavelet bases and obtain conditions of uniform convergence for this wavelets. The next section contains the rate of uniform convergence of Battle-Lemarie wavelet decompositions of stationary Gaussian random processes. Conclusions are made in section 7.

Wavelet Representation of Random Processes

Let $\phi(x),\ x \in \mathbf{R}$ be a function from the space $L_2(\mathbf{R})$ such that $\hat{\phi}(0) \neq 0$ and $\hat{\phi}(y)$ is continuous at 0, where $\hat{\phi}(y) = \int_{\mathbf{R}} e^{-iyx}\phi(x)dx$ is the Fourier transform of ϕ.

Suppose that the following assumption holds true: $\sum_{k\in\mathbf{Z}} |\hat{\phi}(y + 2\pi k)|^2 = 1\ (a.e.)$,

there exists a function $m_0(x) \in L_2([0,2\pi])$, such that $m_0(x)$ has the period 2π and $\hat{\phi}(y) = m_0(y/2)\hat{\phi}(y/2)\ (a.e.)$. In this case the function $\phi(x)$ is called the f-wavelet.

Let $\psi(x)$ be the inverse Fourier transform of the function

$$\hat{\psi}(y) = \overline{m_0\left(\frac{y}{2} + \pi\right)} \cdot \exp\left\{-i\frac{y}{2}\right\} \cdot \hat{\phi}\left(\frac{y}{2}\right).$$

Then the function $\psi(x) = \frac{1}{2\pi}\int_{\mathbf{R}} e^{iyx}\hat{\psi}(y)dy$ is called the m-wavelet.

Let $\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k), \quad \psi_{jk}(x) = 2^{j/2}\psi(2^j x - k), \quad j,k \in \mathbf{Z}.$

It is known that the family of functions $\{\phi_{0k}; \psi_{jk}, j \in \mathbf{N}_0\}$ is an orthonormal basis in $L_2(\mathbf{R})$ (see, for example, [Hardle et al., 1998]).

An arbitrary function $f(x) \in L_2(\mathbf{R})$ can be represented in the form

$$f(x) = \sum_{k\in\mathbf{Z}} \alpha_{0k}\phi_{0k}(x) + \sum_{j=0}^{\infty}\sum_{k\in\mathbf{Z}} \beta_{jk}\psi_{jk}(x), \tag{2}$$

$$\alpha_{0k} = \int_{\mathbf{R}} f(x)\overline{\phi_{0k}(x)}dx, \quad \beta_{jk} = \int_{\mathbf{R}} f(x)\overline{\psi_{jk}(x)}dx.$$

The representation (2) is called a wavelet representation.

The series (2) converges in the space $L_2(\mathbf{R})$ i.e. $\sum_{k\in\mathbf{Z}} |\alpha_{0k}|^2 + \sum_{j=0}^{\infty}\sum_{k\in\mathbf{Z}} |\beta_{jk}|^2 < \infty.$

The integrals α_{0k} and β_{jk} may also exist for functions from $L_p(\mathbf{R})$ and other function spaces. Therefore it is possible to obtain the representation (2) for function classes which are wider than $L_2(\mathbf{R})$.

Let $\{\Omega, B, P\}$ be a standard probability space. Let $\mathbf{X}(t),\ t \in \mathbf{R}$ be a random process such that $\mathbf{EX}(t) = 0$ for all $t \in \mathbf{R}$.

It is possible to obtain representations like (2) for random processes, if their sample trajectories are in the space $L_2(\mathbf{R})$. However the majority of random processes do not possess this property. For example, sample paths of stationary processes are not in $L_2(\mathbf{R})$ (a.s.).

We investigate a representation of the kind (2) for $\mathbf{X}(t)$ with mean-square integrals

$$\xi_{0k} = \int_{\mathbf{R}} \mathbf{X}(t)\overline{\phi_{0k}(t)}dt, \quad \eta_{jk} = \int_{\mathbf{R}} \mathbf{X}(t)\overline{\psi_{jk}(t)}dt.$$

Consider the approximants $\mathbf{X}_{n,\mathbf{k}_n}(t)$ of $\mathbf{X}(t)$ defined by (1).

Assumption S. [Hardle et al., 1998] For the function ϕ there exists a decreasing function $\Phi(x),\ x \ge 0$ such that $\Phi(0) < \infty,\ |\phi(x)| \le \Phi(|x|)$ (a.e.) and $\int_{\mathbf{R}} \Phi(|x|)dx < \infty.$

Let $\mathbf{X}(t)$ be a stationary separable centered Gaussian random process such that its covariance function $R(t,s) = R(t-s)$ is continuous. Let the f-wavelet ϕ and the corresponding m-wavelet ψ be continuous functions and the assumption S holds true for both ϕ and ψ.

Theorem 1 below guarantees the uniform convergence of $\mathbf{X}_{n,\mathbf{k}_n}(t)$ to $\mathbf{X}(t)$.

Theorem 1 [Kozachenko et al., 2011] Suppose that the following conditions hold:

1. There exist $\phi'(u),\ \widehat{\psi}''(u)$, and $\widehat{\psi}(0) = 0,\ \widehat{\psi}'(0) = 0;$

2. $c_\phi := \sup_{u\in\mathbf{R}} |\hat{\phi}(u)| < \infty,\ c_{\phi'} := \sup_{u\in\mathbf{R}} |\hat{\phi}'(u)| < \infty,\ \widehat{\psi}'(u) \in L^1(\mathbf{R}),\ c_{\psi''} := \sup_{u\in\mathbf{R}} |\widehat{\psi}''(u)| < \infty;$

3. $\hat{\phi}(u) \to 0$ and $\widehat{\psi}(u) \to 0$ when $u \to \pm\infty;$

4. There exist $0 < \gamma < \frac{1}{2}$ and $\alpha > \frac{1}{2}$ such that $\int_{\mathbf{R}} \left(\ln(1+|u|)\right)^\alpha |\widehat{\psi}(u)|^\gamma du < \infty,$

$$\int_{\mathbf{R}} \left(\ln(1+|u|)\right)^\alpha |\hat{\phi}(u)|^\gamma du < \infty;$$

5. There exists $\widehat{R}(z)$ and $\sup_{z\in\mathbf{R}} \widehat{R}(z) < \infty;$

6. $\int_{\mathbf{R}} \left|\widehat{R}'(z)\right| dz < \infty$ and $\int_{\mathbf{R}} \left|\widehat{R}^{(p)}(z)\right| |z|^4 dz < \infty$ for $p = 0,1.$

Then $\mathbf{X}_{n,\mathbf{k}_n}(t) \to \mathbf{X}(t)$ uniformly in probability on each interval $[0,T]$ when $n \to \infty$, $k_{0'} \to \infty$ and $k_j \to \infty$ for all $j \in \mathbf{N}_0$.

Conditions of Uniform Convergence for Meyer Wavelets Decompositions of Gaussian Random Processes

Meyer wavelets $\phi(x)$ and $\psi(x)$ cab be given as inverse Fourier transforms of the functions $\hat{\phi}(y)$ and $\widehat{\psi}(y)$ respectively. The expressions of $\hat{\phi}(y)$ and $\widehat{\psi}(y)$ are following:

$$\hat{\phi}(y) = \begin{cases} \frac{1}{\sqrt{2}} \hat{h}(\frac{y}{2}), & |y| \leqslant \frac{4\pi}{3} \\ 0, & |y| > \frac{4\pi}{3} \end{cases} \tag{3}$$

where

$$\hat{h}(y) = \begin{cases} \sqrt{2}, & |y| \leqslant \frac{\pi}{3} \\ 0, & y \in [-\pi, -\frac{2\pi}{3}] \cup [\frac{2\pi}{3}, \pi] \end{cases} \tag{4}$$

and

$$\widehat{\psi}(y) = \begin{cases} 0, & |y| \leqslant \frac{2\pi}{3}, \\ \frac{1}{\sqrt{2}} \widehat{g}(\frac{y}{2}), & \frac{2\pi}{3} \leqslant |y| \leqslant \frac{4\pi}{3} \\ \frac{1}{\sqrt{2}} e^{-i\frac{y}{2}} \hat{h}(\frac{y}{4}), & \frac{4\pi}{3} \leqslant |y| \leqslant \frac{8\pi}{3}, \\ 0, & |y| > \frac{8\pi}{3} \end{cases} \tag{5}$$

where

$$\widehat{g}(y) = e^{-iy} \hat{h}^*(y+\pi).$$

The functions $\phi(x)$ and $\psi(x)$ are C^∞ because their Fourier transforms have a compact support. Wavelet $\psi(x)$ has an infinite number of vanishing moments [Mallat, 1998], so $\widehat{\psi}^{(k)}(0) = 0, k \ge 0.$

Theorem 2 Let $\mathbf{X}(t)$ be a stationary separable centered Gaussian random process such that its covariance function $R(t,s) = R(t-s)$ is continuous. Let ϕ and ψ be Meyer wavelets. Suppose that the following conditions hold:

1. There exists $\widehat{R}(z)$ and $\sup_{z\in\mathbf{R}} \widehat{R}(z) < \infty;$

2. $\int_{\mathbf{R}} \left|\widehat{R}'(z)\right| dz < \infty$ and $\int_{\mathbf{R}} \left|\widehat{R}^{(p)}(z)\right| |z|^4 dz < \infty$ for $p = 0,1.$

Then $\mathbf{X}_{n,\mathbf{k}_n}(t) \to \mathbf{X}(t)$ uniformly in probability on each interval $[0,T]$ when $n \to \infty$, $k_{0'} \to \infty$ and $k_j \to \infty$ for all $j \in \mathbf{N}_0$.

Proof. Statement of this Theorem follows from Theorem 1, if we take into account that assumptions 1) – 4) of Theorem 1 hold for the Meyer wavelets. Indeed, Meyer wavelet $\psi(x)$ has an infinite number of vanishing moments [Mallat, 1998], so $\widehat{\psi}^{(k)}(0) = 0, k \geq 0.$ Now we can use such fact that Fourier transforms of Meyer wavelets have a compact support, so we have fulfillment of conditions 3) and 4) of Theorem 1. Another fact that Fourier transforms of Meyer wavelets is n times continuously differentiable, then assumption 2 of Theorem 1 holds true.

Convergence Rate in the Space $C[0,T]$ of the Meyer Wavelets Representations of Random Processes

In the paper [Kozachenko et al., 2013] an explicit estimate of the rate of uniform convergence for wavelets expansions of Gaussian random processes is obtained. In this section our focus is on the Meyer wavelet bases. So convergence rate in the space $C[0,T]$ for the Meyer wavelets decompositions of Gaussian random processes is studied.

Theorem 3 [Kozachenko et al., 2013] Let $X(t), t \in [0,T]$ be a separable Gaussian stationary random process. Let assumptions of Theorem 1 hold true for $X(t)$.

Then

$$P\left\{\sup_{t\in[0,T]} |\mathbf{X}(t) - \mathbf{X}_{n,\mathbf{k}_n}(t)| > u\right\} \leq 2\exp\left\{-\frac{(u - \sqrt{8u\delta(\varepsilon_{\mathbf{k}_n})})^2}{2\varepsilon_{\mathbf{k}_n}^2}\right\},$$

where $u > 8\delta(\varepsilon_{\mathbf{k}_n})$,

$$\varepsilon_{\mathbf{k}_n} := \sum_{j=0}^{n-1} \frac{A}{2^{j/2}\sqrt{k_j}} + \frac{B}{\sqrt{k_{0'}}} + \frac{C}{2^{n/2}}.$$

A, B, and C are constants which depend only on the covariance function of $\mathbf{X}(t)$ and the wavelet basis. Explicit expressions for A, B, and C are given in the proof of the theorem.

From the paper [Kozachenko et al., 2013] A, B, C, are following:

$$A := B_1^{\psi}\left(6A^{\psi}\sum_{m=1}^{\infty}\frac{1}{m^{3/2}} + 4A_1^{\psi}\right)^{1/2}.$$

$$B := B_1^{\phi}\left(6A^{\phi}\sum_{m=1}^{\infty}\frac{1}{m^{3/2}} + 4A_1^{\phi}\right)^{1/2}.$$

$$C := (2+\sqrt{2})\left(3A^{\psi}\left(B_1^{\psi}\right)^2\left(\sum_{k=1}^{\infty}\frac{1}{k^{\frac{3}{2}}}\right)^2\right.$$

$$+\left(A_1^{\psi}\left(B_1^{\psi}\right)^2 + \frac{c_2 A^{\psi} B_1^{\psi}}{\pi} \right) \sum_{k=1}^{\infty} \frac{1}{k^2} + \frac{c_2^2 A_1^{\psi}}{32\pi^2} \Bigg)^{1/2} .$$

In the case $T = 3$ we can get:

$$B_1^{\phi} = \frac{1}{(2\pi)} \left(\int_{\mathbf{R}} \left| \hat{\phi}'(u) \right| du + T \int_{\mathbf{R}} \left| \hat{\phi}(u) \right| du \right) \approx 3.49516,$$

$$B_1^{\psi} = \frac{1}{(2\pi)} \left(\int_{\mathbf{R}} \left| \widehat{\psi}'(u) \right| du + T \int_{\mathbf{R}} \left| \widehat{\psi}(u) \right| du \right) < 0.01,$$

$$l = \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} = \zeta\left(\frac{3}{2}\right), \ \tilde{l} = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

For the covariance function $R(\tau) = \exp\{-\frac{4}{9}\tau^2\}$ we can obtain the value of the following expressions:

$$A^{\phi} := \frac{1}{2\pi} \left(c_{\phi}^2 \int_{\mathbf{R}} \left| \widehat{R'}(z) \right| dz + 2 c_{\phi} c_{\phi'} \int_{\mathbf{R}} \left| \widehat{R}(z) \right| dz \right) \approx 1.1355,$$

$$A^{\psi} = \frac{c_{\psi''}^2}{2\pi} \left(\int_{\mathbf{R}} | \widehat{R}'(z) || z |^4 dz + 2 \int_{\mathbf{R}} | \widehat{R}(z) || z |^3 dz \right) \approx 3.20112,$$

$$A_1^{\phi} := \frac{c_{\phi}^2}{2\pi} \int_{\mathbf{R}} | \widehat{R}(z) | dz \approx 0.398942,$$

$$A_1^{\psi} := \frac{c_{\psi''}^2}{2\pi} \int_{\mathbf{R}} | \widehat{R}(z) || z |^4 dz \approx 0.945641,$$

$$c_2 := \int_{\mathbf{R}} \left| \widehat{\psi}(v) \right| dv < 0.01.$$

Then we can calculate constants for the expression $\varepsilon_{\mathbf{k}_n}$:

$$A \approx 0.0073456, B \approx 15.3922, C \approx 0.004424.$$

In the paper [Kozachenko et al., 2013] needed formulas for calculation of the following expression are given:

$$\delta(\varepsilon_{\mathbf{k}_n}) := \frac{\gamma}{\sqrt{2}} \left(\sqrt{\ln(T+1)} + \left(1 - \frac{1}{2\alpha}\right)^{-1} \left(\frac{c}{\gamma}\right)^{\frac{1}{2\alpha}} \right),$$

where $\gamma := \min\left(\varepsilon_{\mathbf{k}_n}, \sigma\left(\frac{T}{2}\right) \right), \ \alpha > \frac{1}{2}.$

So for $\alpha = 0.6$ and $\beta = 0.52$ we can evaluate following constants:

$$c = B_0 + B_1 + B_2$$

$$B_0 := \left(q_1 + A^{\psi} Q_1 K^2 \right)^{1/2} \cdot \sum_{j=0}^{\infty} \frac{(j+1)^{\alpha}}{2^{j/2}} \approx 0.107345,$$

$$B_1 := \left(q + q_1 + q_2 + A^{\psi} Q K^2\right)^{1/2} \cdot \sum_{j=0}^{\infty} \frac{(j+1)^{\alpha}}{2^{j/2}} \approx 0.108815,$$

$$B_2 := \left(q_{\phi 1} + A^{\phi} (K^{\phi})^2 Q\right)^{1/2} \approx 1308.01.$$

$$K := \pi^{-1} \Big(2^{3+\alpha-\beta} \pi^{\beta} c_{\psi'}^{\beta} \left((\ln 5)^{\alpha} c_0 + c_1 \right) +$$

$$+ \pi T 2^{\alpha-1} \left((\ln 5)^{\alpha} c_2 + c_3 \right) + c_{\alpha} c_2 \Big) < 0.01.$$

$$c_0 := \int_{\mathbf{R}} \left|\widehat{\psi}(v)\right|^{1-\beta} dv < \infty, \quad c_1 := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\widehat{\psi}(v)\right|^{1-\beta} dv < \infty.$$

$$c_2 := \int_{\mathbf{R}} \left|\widehat{\psi}(v)\right| dv < \infty, \quad c_3 := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\widehat{\psi}(v)\right| dv < \infty.$$

$$Q \le Q_1 = \left(\sum_{k=1}^{\infty} \frac{1}{2k^{\frac{1}{2}+\beta}} \right)^2 + c_{\delta}^{\beta} \sum_{m=1}^{\infty} \frac{1}{m^{1+\delta\beta}} \sum_{l=1}^{\infty} \frac{1}{l^{(2-\delta)\beta}} \approx 111.259,$$

$$q := \frac{2^{\alpha} A^{\psi} K((\ln 5)^{\alpha} c_2 + c_3)}{\pi} \cdot \sum_{l=1}^{\infty} \frac{1}{l^{1+\beta}} \approx 8.8 \times 10^{-6},$$

$$q_1 := \frac{A_1^{\psi} K^2}{2} \cdot \sum_{k=1}^{\infty} \frac{1}{k^{2\beta}} \approx 5.0 \times 10^{-6},$$

$$q_2 := \frac{2^{2\alpha} A_1^{\psi}}{\pi^2} \left((\ln 5)^{\alpha} c_2 + c_3 \right)^2 \approx 1.2 \times 10^{-6},$$

$$q_{\phi 1} := \frac{A_1^{\phi} (K^{\phi})^2}{2} \cdot \sum_{k=1}^{\infty} \frac{1}{k^{2\beta}} \approx 8348.11.$$

$$K^{\phi} := \pi^{-1} \Big(2^{3+\alpha-\beta} \pi^{\beta} c_{\phi'}^{\beta} \left((\ln 5)^{\alpha} c_{\phi 0} + c_{\phi 1} \right) +$$

$$\pi T 2^{\alpha-1} \left((\ln 5)^{\alpha} c_{\phi 2} + c_{\phi 3} \right) + c_{\alpha} c_{\phi 2} \Big) \approx 116.087,$$

$$c_{\phi 0} := \int_{\mathbf{R}} \left|\hat{\phi}(v)\right|^{1-\beta} dv < \infty, \quad c_{\phi 1} := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\hat{\phi}(v)\right|^{1-\beta} dv < \infty,$$

$$c_{\phi 2} := \int_{\mathbf{R}} \left|\hat{\phi}(v)\right| dv < \infty, \quad c_{\phi 3} := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\hat{\phi}(v)\right| dv < \infty.$$

So, if we take into consideration this calculation, then c is following:

$$c \approx 1308.22.$$

Naturally:

$$\sigma(T) = \frac{c}{\left(\ln\left(e^{\alpha} + \frac{1}{T} \right) \right)^{\alpha}} \approx 1532.73.$$

$X_{n,k_j}(t)$ approximates process $X(t)$ with the reliability of $1-\tilde{\delta}$ and accuracy $\tilde{\varepsilon}$, if

$$\mathbf{P}\left\{\sup_{0\le t\le T} | X(t) - X_{n,k_j}(t) | > \tilde{\varepsilon}\right\} \le \tilde{\delta}.$$

Let $\tilde{\delta} = 0.01$, we can use the rule of three σ, then it can be considered as $\tilde{\varepsilon} = 0.1 \cdot 6\sigma$. In our case, for the covariance function $R(\tau) = \exp\{-\frac{4}{9}\tau^2\}$, we can calculate the variance $\sigma = 1$, so $\tilde{\varepsilon} = 0.6$. Then, for Meyer wavelets, using the program Wolfram Mathematica, we can obtain such $\tilde{\delta} = 0.01$ at $k_0 = 85$, $k_j = 20$, $n = 20$ with a slight increase k_0, $\tilde{\delta}$ is significantly reduced.

Conditions of Uniform Convergence for Battle-Lemarie Wavelets Expansions of Gaussian Random Processes

Polynomial spline wavelets introduced by Battle and Lemarie are computed from spline multiresolution approximations. Let $\phi_m(x)$ and $\psi_m(x)$ be the inverse Fourier transforms of the functions $\hat{\phi}_m(y)$ and $\widehat{\psi}_m(y)$ respectively. The expressions of $\hat{\phi}_m(y)$ and $\widehat{\psi}_m(y)$ are following:

$$\hat{\phi}_m(y) = \frac{e^{-i\frac{y\cdot\varepsilon}{2}}}{y^{m+1}\sqrt{S_{2m+2}(y)}}, \tag{6}$$

where

$$S_n(y) = \sum_{k=-\infty}^{\infty} \frac{1}{(y+2k\pi)^n} \tag{7}$$

and $\varepsilon = 1$ if m is even and $\varepsilon = 0$ if m odd.

$$\widehat{\psi}_m(y) = \frac{e^{-i\frac{y}{2}}}{y^{m+1}}\sqrt{\frac{S_{2m+2}(\frac{y}{2}+\pi)}{S_{2m+2}(y)S_{2m+2}(\frac{y}{2})}}. \tag{8}$$

For the $m-$ degree spline wavelet $\psi(x)$ has $m+1$ vanishing moments [Mallat, 1998], so $\widehat{\psi}^{(k)}(0) = 0, 0 \le k \le m+1.$ Wavelet $\psi(x)$ has an exponential decay. Since it is a polynomial spline of degree $m,$ it is $m-1$ times continuously differentiable (see, for example, [Mallat, 1998]).

To check the assumption 1 of Theorem 1 we can use Lemma 1 from the paper [Polosmak, 2009]:

Lemma 1 [Polosmak, 2009] Let $\phi(x)$ - such function, that $\int_{\mathbf{R}} |\phi(x)|dx < \infty, \phi(x) \to 0, x \to \pm\infty,$ let the derivative $\phi'(x)$ exists, such that $\int_{\mathbf{R}} |\phi'(x)|^{\gamma} dx < \infty$ for some $0 < \gamma < 1.$ Let $|\phi'(x) - \phi'(y)| \le \sigma(|x-y|),$ where $\sigma = \{\sigma(u), u > 0\}$ such monotone increasing function that $\sigma(0) = 0.$

Then $\int_{\mathbf{R}} |\hat{\phi}(y)| c(y)dy < \infty,$ where $\hat{\phi}(y) = \int_{\mathbf{R}} e^{-iyx}\phi(x)dx,$ and $c = \{c(y), y \in \mathbf{R}\},$ $c(y) > 0$ such function, that $\int_1^{\infty} \frac{1}{|y|}\left(\sigma\left(\frac{\pi}{y}\right)\right)^{1-\gamma} c(y)dy < \infty.$

Remark 1 In the case of $\sigma(u) = c\,|u|^{\alpha}$, $0 < \alpha \le 1$, we can take $c(y) = \ln(1+|y|)^{\delta}$, where $\delta > \frac{1}{2}$.

Theorem 4 Let $\mathbf{X}(t)$ be a stationary separable centered Gaussian random process such that its covariance function $R(t,s) = R(t-s)$ is continuous. Let ϕ and ψ be Battle-Lemarie wavelets. Suppose that the following conditions hold:

1. There exists $\widehat{R}(z)$ and $\sup_{z\in\mathbf{R}} \widehat{R}(z) < \infty$;

2. $\int_{\mathbf{R}} \left|\widehat{R}'(z)\right| dz < \infty$ and $\int_{\mathbf{R}} \left|\widehat{R}^{(p)}(z)\right| |z|^4 dz < \infty$ for $p = 0,1$.

Then $\mathbf{X}_{n,\mathbf{k}_n}(t) \to \mathbf{X}(t)$ uniformly in probability on each interval $[0,T]$ when $n \to \infty$, $k_{0'} \to \infty$ and $k_j \to \infty$ for all $j \in \mathbf{N}_0$.

Proof. Statement of this Theorem follows from Theorem 1, if we take into account that assumptions 1) – 4) of Theorem 1 hold for the Battle-Lemarie wavelets. Indeed, $m-$ degree ψ - wavelet Battle-Lemarie has $m+1$ vanishing moments [Mallat, 1998], so $\widehat{\psi}^{(k)}(0) = 0, 0 \le k \le m+1$. Another fact that it is $m-1$ times continuously differentiable, then, using formulas (3),(5), we have fulfillment of conditions 1) – 3) of Theorem 1. Assumption 4 follows from Lemma 1, differentiability of the Battle-Lemarie wavelets and Remark 1.

Convergence Rate in the Space $C[0,T]$ of the Battle-Lemarie Wavelets Representations of Random Processes

In the previous section it was given Theorem 3 from the paper [Kozachenko et al., 2013] in which an explicit estimate of the rate of uniform convergence for wavelets expansions of Gaussian random processes is obtained. In this section our focus is on the Battle-Lemarie wavelet bases. So convergence rate in the space $C[0,T]$ for the Battle-Lemarie wavelets decompositions of Gaussian random processes is studied. Here we obtain all constants for Theorem 3 in the case of Battle-Lemarie wavelets:

$$A := B_1^{\psi}\left(6A^{\psi}\sum_{m=1}^{\infty}\frac{1}{m^{3/2}} + 4A_1^{\psi}\right)^{1/2}.$$

$$B := B_1^{\phi}\left(6A^{\phi}\sum_{m=1}^{\infty}\frac{1}{m^{3/2}} + 4A_1^{\phi}\right)^{1/2}.$$

$$C := (2+\sqrt{2})\left(3A^{\psi}\left(B_1^{\psi}\right)^2\left(\sum_{k=1}^{\infty}\frac{1}{k^{\frac{3}{2}}}\right)^2 + \left(A_1^{\psi}\left(B_1^{\psi}\right)^2 + \frac{c_2 A^{\psi} B_1^{\psi}}{\pi}\right)\sum_{k=1}^{\infty}\frac{1}{k^2} + \frac{c_2^2 A_1^{\psi}}{32\pi^2}\right)^{1/2}.$$

In the case $T = 3$ we can get:

$$B_1^{\phi} = \frac{1}{(2\pi)}\left(\int_{\mathbf{R}}\left|\hat{\phi}(u)\right| du + T\int_{\mathbf{R}}\left|\hat{\phi}(u)\right| du\right) \approx 3.97712,$$

$$B_1^{\psi} = \frac{1}{(2\pi)}\left(\int_{\mathbf{R}}\left|\widehat{\psi}'(u)\right|du + T\int_{\mathbf{R}}\left|\widehat{\psi}(u)\right|du\right) \approx 4.64062,$$

$$l = \sum_{k=1}^{\infty}\frac{1}{k^{3/2}} = \zeta\left(\frac{3}{2}\right),\ \tilde{l} = \sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}.$$

For the covariance function $R(\tau) = \exp\{-\frac{4}{9}\tau^2\}$ we can obtain the value of the following expressions:

$$A^{\phi} := \frac{1}{2\pi}\left(c_{\phi}^2\int_{\mathbf{R}}\left|\widehat{R}'(z)\right|dz + 2c_{\phi}c_{\phi'}\int_{\mathbf{R}}\left|\widehat{R}(z)\right|dz\right) \approx 1.1355,$$

$$A^{\psi} = \frac{c_{\psi''}^2}{2\pi}\left(\int_{\mathbf{R}}|\widehat{R}'(z)||z|^4 dz + 2\int_{\mathbf{R}}|\widehat{R}(z)||z|^3 dz\right) \approx 3.20112,$$

$$A_1^{\phi} := \frac{c_{\phi}^2}{2\pi}\int_{\mathbf{R}}|\widehat{R}(z)|dz \approx 0.398942,$$

$$A_1^{\psi} := \frac{c_{\psi''}^2}{2\pi}\int_{\mathbf{R}}|\widehat{R}(z)||z|^4 dz \approx 0.945641,$$

$$c_2 := \int_{\mathbf{R}}\left|\widehat{\psi}(v)\right|dv \approx 7.5725.$$

Then we can calculate constants for the expression $\varepsilon_{\mathbf{k}_n}$:

$$A \approx 0.0073456,\ B \approx 17.5147,\ C \approx 129.78.$$

In the paper [Kozachenko et al., 2013] needed formulas for calculation of the following expression are given:

$$\delta(\varepsilon_{\mathbf{k}_n}) := \frac{\gamma}{\sqrt{2}}\left(\sqrt{\ln(T+1)} + \left(1-\frac{1}{2\alpha}\right)^{-1}\left(\frac{c}{\gamma}\right)^{\frac{1}{2\alpha}}\right),$$

where $\gamma := \min\left(\varepsilon_{\mathbf{k}_n}, \sigma\left(\frac{T}{2}\right)\right),\ \alpha > \frac{1}{2}.$

So for $\alpha = 0.6$ and $\beta = 0.52$ we can evaluate following constants:

$$c = B_0 + B_1 + B_2$$

$$B_0 := \left(q_1 + A^{\psi}Q_1K^2\right)^{1/2}\cdot\sum_{j=0}^{\infty}\frac{(j+1)^{\alpha}}{2^{j/2}} \approx 0.107345,$$

$$B_1 := \left(q + q_1 + q_2 + A^{\psi}QK^2\right)^{1/2}\cdot\sum_{j=0}^{\infty}\frac{(j+1)^{\alpha}}{2^{j/2}} \approx 0.108815,$$

$$B_2 := \left(q_{\phi 1} + A^{\phi}(K^{\phi})^2 Q\right)^{1/2} \approx 1416.2.$$

$$K := \pi^{-1}\left(2^{3+\alpha-\beta}\pi^{\beta}c_{\psi'}^{\beta}\left((\ln 5)^{\alpha}c_0 + c_1\right) +\right.$$

$$\left. + \pi T 2^{\alpha-1}\left((\ln 5)^{\alpha}c_2 + c_3\right) + c_{\alpha}c_2\right) \approx 191.063.$$

$$c_0 := \int_{\mathbf{R}} \left|\widehat{\psi}(v)\right|^{1-\beta} dv < \infty, \quad c_1 := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\widehat{\psi}(v)\right|^{1-\beta} dv < \infty.$$

$$c_2 := \int_{\mathbf{R}} \left|\widehat{\psi}(v)\right| dv < \infty, \quad c_3 := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\widehat{\psi}(v)\right| dv < \infty.$$

$$Q \le Q_1 = \left(\sum_{k=1}^{\infty} \frac{1}{2k^{\frac{1}{2}+\beta}}\right)^2 + c_\delta^\beta \sum_{m=1}^{\infty} \frac{1}{m^{1+\delta\beta}} \sum_{l=1}^{\infty} \frac{1}{l^{(2-\delta)\beta}} \approx 111.259,$$

$$q := \frac{2^\alpha A^\psi K((\ln 5)^\alpha c_2 + c_3)}{\pi} \cdot \sum_{l=1}^{\infty} \frac{1}{l^{1+\beta}} \approx 8.8 \times 10^{-6},$$

$$q_1 := \frac{A_1^\psi K^2}{2} \cdot \sum_{k=1}^{\infty} \frac{1}{k^{2\beta}} \approx 5.0 \times 10^{-6},$$

$$q_2 := \frac{2^{2\alpha} A_1^\psi}{\pi^2} \left((\ln 5)^\alpha c_2 + c_3\right)^2 \approx 1.2 \times 10^{-6},$$

$$q_{\phi 1} := \frac{A_1^\phi (K^\phi)^2}{2} \cdot \sum_{k=1}^{\infty} \frac{1}{k^{2\beta}} \approx 9786.24.$$

$$K^\phi := \pi^{-1}\left(2^{3+\alpha-\beta} \pi^\beta c_{\phi'}^\beta \left((\ln 5)^\alpha c_{\phi 0} + c_{\phi 1}\right) + \right.$$

$$\left. \pi T 2^{\alpha-1} \left((\ln 5)^\alpha c_{\phi 2} + c_{\phi 3}\right) + c_\alpha c_{\phi 2}\right) \approx 125.689,$$

$$c_{\phi 0} := \int_{\mathbf{R}} \left|\hat{\phi}(v)\right|^{1-\beta} dv < \infty, \quad c_{\phi 1} := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\hat{\phi}(v)\right|^{1-\beta} dv < \infty,$$

$$c_{\phi 2} := \int_{\mathbf{R}} \left|\hat{\phi}(v)\right| dv < \infty, \quad c_{\phi 3} := \int_{\mathbf{R}} \left(\ln(1+|v|)\right)^{\alpha} \left|\hat{\phi}(v)\right| dv < \infty.$$

So, if we take into consideration this calculation, then c is following:

$$c \approx 1416.41.$$

Naturally:

$$\sigma(T) = \frac{c}{\left(\ln\left(e^\alpha + \frac{1}{T}\right)\right)^\alpha} \approx 1659.48.$$

$X_{n,k_j}(t)$ approximates process $X(t)$ with the reliability of $1-\tilde{\delta}$ and accuracy $\tilde{\varepsilon}$, if

$$\mathbf{P}\left\{\sup_{0\le t\le T} |X(t) - X_{n,k_j}(t)| > \tilde{\varepsilon}\right\} \le \tilde{\delta}.$$

Let $\tilde{\delta} = 0.01$, we can use the rule of three σ, then it can be considered as $\tilde{\varepsilon} = 0.1 \cdot 6\sigma$. In our case, for the covariance function $R(\tau) = \exp\{-\frac{4}{9}\tau^2\}$, we can calculate the variance $\sigma = 1$, so $\tilde{\varepsilon} = 0.6.$ Then, for Battle-Lemarie wavelets, using the program Wolfram Mathematica, we can obtain such $\tilde{\delta} = 0.01$ at $k_0 = 110$, $k_j = 20$, $n = 20$ with a slight increase k_0, $\tilde{\delta}$ is significantly reduced.

Conclusion

Conditions of uniform convergence for Meyer wavelet decompositions and Battle-Lemarie wavelet decompositions of stationary Gaussian random processes are presented. The rate of convergence in the space $C([0,T])$ of Meyer wavelet decompositions and Battle-Lemarie wavelet decompositions of stationary Gaussian random processes are obtained. We can conclude that both wavelet bases are good for expansion of stationary Gaussian random processes, but the Meyer wavelets have some advantages. For the same accuracy of the approximation in the case of the Meyer wavelets, we need fewer terms in the expansion.

Bibliography

[Atto et al., 2010] A. Atto. Wavelet packets of nonstationary random processes: contributing factors for stationarity and decorrelation. A. Atto, Y. Berthoumieu. IEEE Trans Inform Theory 58(1), p. 317–330, 2010.

[Bardet et al., 2010] J.M. Bardet. A wavelet analysis of the Rosenblatt process: chaos expansion and estimation of the self-similarity parameter. J.M. Bardet, C.A. Tudor. Stochastic Process Appl, 120(12), p. 2331–2362, 2010.

[Buldygin et al., 2000] V. V. Buldygin. Metric characterization of random variables and random processes. V. V. Buldygin and Yu. V. Kozachenko. American Mathematical Society, Providence RI, 2000.

[Daubechies, 1992] I. Daubechies. Ten lectures on wavelets. Society for industrial and applied mathematics, Philadelphia, 1992.

[Didier et al., 2008] G. Didier. Gaussian stationary processes: adaptive wavelet decompositions, discrete approximations and their convergence. G. Didier, V. Pipiras. J Fourier Anal and Appl, 14, p. 203-234, 2008.

[Hardle et al., 1998] W. Hardle Wavelet, Approximation and statistical applications. W. Hardle, G. Kerkyacharian, D. Picard, A. Tsybakov. Springer, New York, 1998.

[Kozachenko et al., 2008] Yu. V. Kozachenko. Uniform convergence in probability of wavelet expansions of random processes from $L_2(\Omega)$. Yu. V. Kozachenko, O. V. Polosmak. Random operators and stochastic equations. 16 (4), p. 12–37, 2008.

[Kozachenko et al., 2011] Yu. V. Kozachenko. Uniform convergence of wavelet expansions of Gaussian random processes. Yu. V. Kozachenko, A. Ya. Olenko, O. V. Polosmak. Stoch. Anal. Appl. 29(2), p. 169–184, 2011.

[Kozachenko et al., 2013] Yu. V. Kozachenko. Convergence rate of wavelet expansions of Gaussian random processes, Comm. Statist. Theory Methods 42, p. 3853–3872, 2013.

[Kozachenko, 2004] Yu. V. Kozachenko. Lectures on wavelet analysis. (Ukrainian), TViMC Kyiv, 2004.

[Mallat, 1998] Stephane Mallat. A wavelet tour of signal processing. Academic Press, New York, 1998.

[Polosmak, 2009] O. V. Polosmak. Rate of uniform convergence for expansions of random processes from $L_2(\Omega)$ on Battle-Lemarie wavelets. Bulletin of University of Kyiv. Series: Physics and Mathematics, 3, p. 36–41, 2009.

Authors' Information

Olga Polosmak *– PhD, assistant of economic cybernetics department, faculty of economics, Taras Shevchenko national university of Kiev, Kiev, Ukraine, olgapolosmak@yandex.ru*

Major Fields of Scientific Research: Random processes, Wavelet analysis, Wavelet expansions of stochastic

WIRELESS DATA TRANSMISSION OPTIONS IN ROTARY IN-DRILLING ALIGNMENT(R-IDA) SETUPS FOR MULTILATERAL OIL DRILLING APPLICATIONS

Zhenhua Wang, Tao Li, Myles McDougall, Dan McCormack, Martin P. Mintchev

Abstract: *Conventional methods in multilateral drilling processes incorporate magnetometer-based surveying systems for determining the position and attitude of the bottomhole assembly (BHA). Magnetic surveying results in an increased weight of the BHA, higher cost due to shielding with bulky nonmagnetic collars and, more importantly, severely degraded performance due to unavoidable geomagnetic interferences such as metal and ore deposits in the vicinity. Micro-Electromechanical Systems (MEMS) based Inertial Navigation Systems (INS) have been proposed as an alternative to magnetic surveying for multilateral drilling. Previous studies have shown theoretically and experimentally that a Kalman filter-based In-Drilling Alignment (IDA) and its minimized version Rotary In-Drilling Alignment (R-IDA) successfully limit the accumulated error growth associated with the INS, compared to the traditional zero-velocity update (ZUPT) alignment method. A high performance wireless MEMS-based INS has been proposed for R-IDA alignment, the implementation simplicity and high accuracy of which could be very useful in multilateral drilling processes. Furthermore, this paper discusses the concept of wireless data transmission within drill pipes downhole in comparison to other existing or emerging methods. It is shown that wireless telemetry inside the drill pipes is potentially capable of transmitting up to 250,000 bits per second (bits/sec) with high reliability and low power consumption, which makes a drill string based local network for real-time downhole monitoring and control applications feasible.*

Keywords: *Multilateral drilling design, wireless data transmission, downhole instrumentation, rotary In-Drilling Alignment.*

ACM Classification Keywords: *A.0 General Literature - Conference proceedings; J.2 Physical Sciences and Engineering.*

Introduction

A. Multilateral oil drilling applications

Multilateral drilling technology is a recently emerging evolution in oil well exploration in which several smaller wellbore branches deviate from their parent borehole [Pasicznyk, 2001] (Figure 1a). Horizontal drilling (HD) has a better chance of intersecting more fractures than a vertical well, but there is a limit to how far horizontal wells can be drilled. By drilling other laterals from the same wellbore, twice the number of fractures can often be exposed at a much lower cost than drilling long horizontal sections or another well [Bosworth et al., 1998]. Horizontal wells and their related branches usually target the same reservoir interval. The goal of this type of well is to increase production rates, improve hydrocarbon recovery and maximize production from that zone.

Directional drilling techniques can be used to drill vertical multilateral wells, which is the other type of multilateral drilling. Vertical multilaterals are additional wells sunk down from a main wellbore (Figure 1b) and stretched laterally to up to 100 meters. Multilaterals can be as simple as an open hole sidetrack, or they can be more sophisticated with a junction that is cased and has pressure isolation and reentry capabilities [Bosworth et al, 1998]. Multilaterals are usually being used where production can be incrementally increased with less capital costs. They can also be employed for offshore drilling where the numbers of slots are limited [Bosworth et al, 1998].

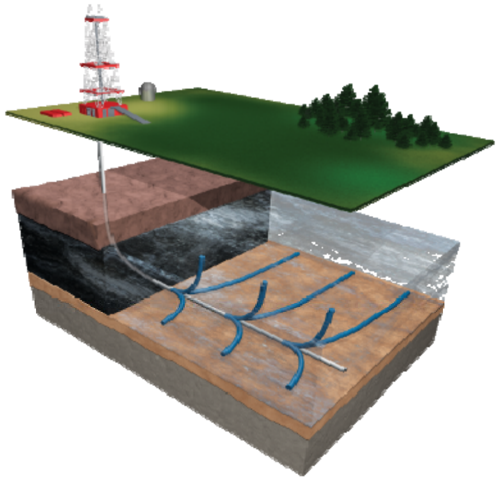


Figure 1a. Horizontal multilateral oil drilling

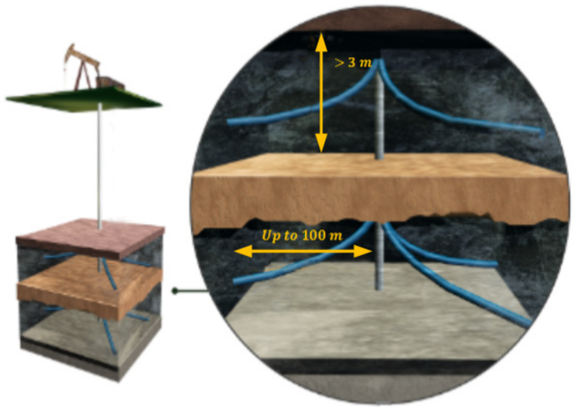


Figure 1b. Vertical multilateral oil drilling

A successful multilateral well that replaces several vertical wellbores can reduce overall drilling and completion costs, increase production and provide more efficient drainage of a reservoir. Furthermore, multilaterals can make reservoir management more efficient and help increase recoverable reserves.

B. Rotary In-Drilling Alignment (R-IDA) for error compensation in downhole navigation systems

Traditional downhole navigation is based on the so-called Measurement-While-Drilling (MWD) systems, which incorporate triad orthogonal accelerometers and triad orthogonal magnetometers to determine the position and altitude of the bottomhole assembly (BHA). However, the magnetometers are susceptible to external magnetic interferences, including randomly located ore deposits, drill string rotations and drilling fluid circulations [Shelkholeslami et al, 1991; Torkildsen et al, 2004]. Typically, magnetic interferences are addressed by shielding the BHA with very expensive and bulky nonmagnetic collars, which can only minimize, rather than eliminate magnetic artifacts. Degradation of the magnetometers is sometimes dramatic and the oil industry has to employ alternative wireline gyroscopes to determine BHA's attitude, which costs it tremendous time and money.

An inertial navigation system (INS) is autonomous dead-reckoning (DR) approach of tracking the position and orientation of an object by 3 mutually perpendicular accelerometers and 3 mutually perpendicular gyroscopes, which are not influenced by magnetic interferences. Commercially available Micro-Electromechanical Systems (MEMS)-based inertial measurement units (IMU) that contain accelerometers and gyroscopes are excellent candidates for complete INS due to their low cost, small size and low power consumption, which are critical factors for downhole drilling. However, the unlimited error growth in the measurements might be prohibitive for the long-term utilization of this technology. New methods to reduce the INS error include In-Drilling Alignment (IDA) [Jurkov et al, 2011] and its reduced version, Rotary In-Drilling Alignment (R-IDA) [Wang et al, 2013], which are both utilized to increase the observability for their Kalman filters (KF) by precisely inducing controlled motions during operational breaks in the drilling process. Previous study [Wang et al, 2013] has shown that a MEMS-based autonomous IMU produced by Memsense LLC (Rapid City, SD, USA) employing the R-IDA method achieved error reduction 2 times greater than the same device utilizing the mainstream error compensation method known as zero-velocity update (ZUPT). However, these results were obtained in laboratory conditions, and the autonomous Memsense IMU noise and temperature characteristics were not adequate for real downhole oil drilling applications.

C. Problems of data transmission in downhole navigation

Harsh downhole environment encountered in drilling applications presents particular challenges, especially for reliable data transmission between the surface and the drilling devices due to the high pressure, extreme temperatures and huge distances downhole. At the bottom of the drill string is the BHA, which includes the drill bit along with electronic components such as sensors, control mechanisms and required circuitry [PetroWiki, 2013]. The sensors in the BHA typically contain surveying of various properties of the formation and the fluid within it as well as the navigation measurements. The processes that require information from such downhole sensors include MWD and Logging-While-Drilling (LWD) systems. Once MWD/LWD logs the downhole parameters and measurements, there are mainly three methods to pull the information upward. They are discussed separately below.

Mud-pulse telemetry (MPT)

Mud-pulse telemetry (MPT) is the most common and standard method of data transmission downhole and is especially used by MWD/LWD tools developed in the 1970's [Wasserman et al., 2008]. The conventional MWD/LWD tool incorporates an electronic sensors package and a mudflow wellbore telemetry device. The drilling fluid called mud is pumped from the surface to the downhole BHA along the drill string. The mud serves as a cooling and lubricating circulation fluid and continuously carries the derbies back up to the ground during the drilling process. The mudflow wellbore telemetry device can selectively restrict the passages of the mud through the drill string to control and manipulate the pressure in the mud lines by operating a valve. These manipulations create pressure fluctuations which represent information that is being encoded in binary format and propagated within the mud towards the surface where it is received from pressure transducers [Wasserman et al., 2008]. The problems with this type of data transmission are exactly two: (1) slow speed of the serial mud-based interface; and (2) high power consumption.

Electromagnetic telemetry (EMT)

Electromagnetic telemetry (EMT) employs a downhole current source to emit an electromagnetic signal into a formation. The signal can be detected and received at the surface due to a small voltage drop between the top part of the BHA (the main drill string) and the bottom part of the BHA. Typically, the EM tool can generate voltage difference between the drill string sections at a very low frequency, below 30 Hz [Gao et al, 2006]. The information then is converted into modulated EM waves by digital modulation. The typical transmission rate of the EMT tool is around 10 bits/sec. Compared to the MPT, the EMT method does not require changes in major drilling parameters such as rotation of the drill pipe and mud flow rate, to send information to the surface. It also does not rely on the composition of the mud flow, since most of the mud is compressed, gas-filled fluid for the underbalanced drilling (UBD) in order to reduce the equivalent density. This results in high signal attenuation during the data transmission that severely handicaps the MPT's communication capability. EMT tool is usually employed for certain specific applications such as UBD because of its immunity to the drilling fluid. However, EMT can also lose strength dramatically in some types of formations, becoming unpredictable at several thousand feet of depth [Gao et al., 2006]. Its cost is also significant.

Wired cable data transmission

Some research has been focusing on the development of wired drill pipe system since the beginning of the 21st century [National Oilwell Varco, 2014; Jellison et al., 2003]. It is based on the theory that composite drill pipes can also facilitate high-speed data transfer rates via special materials such as fiber optic cables embedded within the pipes during construction. A great benefit of such system is its superior data transmission rate which makes a real-time monitoring system downhole in drilling completely feasible. The IntelliServ wired pipe, offering data rates upwards of 1 million bits/sec became commercial in 2006 [National Oilwell Varco, 2014]. However, cables can cause reliability problems attributed to installation, connections and drilling fluid migration. In addition, the high capital costs and lack of advanced technology in drill pipe material and manufacturing is still limiting its development.

Since the data transmission rate of the MPT and EMT correlated with bandwidth is less than 100 bits/sec [Jellison et al, 2003], most of the useful information provided by the MWD and LWD will be lost or stored in a memory logger associated with the downhole instrumentation near the drill bit. Therefore, the so-called "real-time drilling" can be applied in a very limited scope. Comparisons of the commercialized downhole transmission methods are given below (Table 1).

Table 1. Comparison between different downhole data transmission methods

Transmission Method	MPT	EMT	Wired Cable Drill Pipe
Data Bandwidth [Bits/sec]	1-40 [Wasserman et al., 2008]	10-100 [Jellison et al., 2003]	1,000,000 [National Oilwell Varco, 2014]
Reliability	Low	High	Low
Applications	Limited to UBD	Susceptible to formation	Wide
Working Depth [Meters]	Up to 12,000	Up to 3,000 [Baker Hughes, 2014]	Up to 15,000
Frequency [Hz]	<100	<30	N/A
Cost	Low	Medium	Very High

D. Aim of the paper

This article aims at proposing a specific R-IDA design concept for multilateral drilling applications, including an efficient wireless communication for the entire setup. It also reviews some of the competitive existing, emerging or theoretically-discussed approaches and their actual and potential drawbacks.

Methods

A. Rotary In-Drilling Alignment (R-IDA) design for multilateral drilling navigation

For lateral drilling applications, a capsule-based, high-performance IMU wireless module has been proposed in R-IDA context. The capsule includes microcontroller unit (MCU), wireless radio frequency (RF) module and IMU, all of which are finally integrated on a printed circuit board (PCB) to be packaged for downhole mounting. The stepper motor-based system [Wang et al, 2013] rotates the capsule freely at a certain speed during the scheduled operational stops. The data from the wireless IMU module are transmitted to an embedded wireless RF module residing at an appropriate distance. The RF module then automatically sends the received data to another module. By setting a number of wireless RF modules within the entire drill string at certain distances from each other, the wireless IMU measurements are transmitted along the drill string towards the surface. A computer equipped with a receiver at the surface captures and decodes downhole data and runs the Kalman filtering (KF)-based navigation algorithms [Wang et al, 2013] to compute attitude and position of the BHA within a very short delay.

B. Downhole wireless transmission options (Pros and Cons)

A wireless communication system driven by a downhole battery could allow intelligent transmission sensors to be placed anywhere, avoiding the need for cables to supply power or to transmit data. Currently, the oil industry does not employ mature wireless communication systems downhole. However, more research concentrates on this area due to its huge potential benefits. The concept of acoustic wireless communication has existed for several decades and it came to the verge of being commercialized. Wireless acoustic telemetry, based on the propagation of stress waves along the drill pipe, requires less power than conventional systems such as the EMT and MPT [Kyle et al, 2013]. It also shows a potential at higher data transmission rate capability of 50-100 bits/sec through the drill pipe channel [Gao et al, 2006]. However, it is still of relatively low frequency (400-2000Hz) and slow transmission rate, which is not applicable for real time monitoring. Furthermore, it is very susceptible to drill string interferences, since it is an acoustic transmission channel, where many passbands and stopbands occur [Gao et al, 2006]. As a result, reflected and transmitted acoustic signals interfere to the point where they are totally suppressed.

Other commonly used standard wireless RF communication methods including the WiFi (IEEE 802.11) and Bluetooth (IEEE 802.15.1) have unprecedented transmission rate operating at 2.4 GHz frequency. However, they are not applicable downhole due to the power consumption, reliability, and complexity etc. [Farahani, 2011].

Results

A. Proposed Rotary In-Drilling Alignment (R-IDA) setup

A high performance MEMS-based IMU ADIS16488A (Analog Devices Inc., Norwood, MA, USA) was selected for the proposed R-IDA design. This IMU provides tactical grade precision of the gyroscope measurements. The angular velocity range is 450°/s with a gyroscope bias instability of 5.1°/h and angle random walk (ARW) of $0.26°/\sqrt{h}$. The system is very compact at 47 × 44 × 14 mm, 48 g weight, and uses 0.8 W of power [Analog Devices, 2014]. The temperature resistance standalone is as high as 105 degrees Celsius, which in conjunction with its low power consumption and small size, presents it as a very competent candidate for multilateral drilling applications. However, the ADIS16488A has a wired cable connection interface based on the Serial Peripheral Interface (SPI) protocol, which is prohibitive for R-IDA applications. At the IMU end, it is proposed to employ an XBee Pro Series 1 (Digi International Inc., MN, USA) RF wireless module. The selected wireless module is ZigBee (IEEE 802.15.4)-based protocol with an asynchronous interface while the IMU's interface is synchronous. To allow these two different interfaces to communicate in real time, a low-power mbed microcontroller LPC11u24 (ARM Holdings plc, England) with an ARM-Cortex M0 processor embedded is incorporated. The role of the microcontroller is to perform an SPI-to-ZigBee interface translation operation since the ADIS16488A's interface is SPI-based while the XBee Pro S1 wireless module is UART-based. At the computer end, free open source Tera Term software (Tera Term Project, Japan) was installed to display and log the IMU measurements in real time through the wired connection with the XBee Explorer USB (Sparkfun Electronics Inc., CO, USA), which acts as a receiver. The block diagram in Figure 2 illustrates the high-level wireless design suggested in this paper. The real wireless navigation module (Figure 3) will also incorporate a downhole battery on a PCB in a casing package.

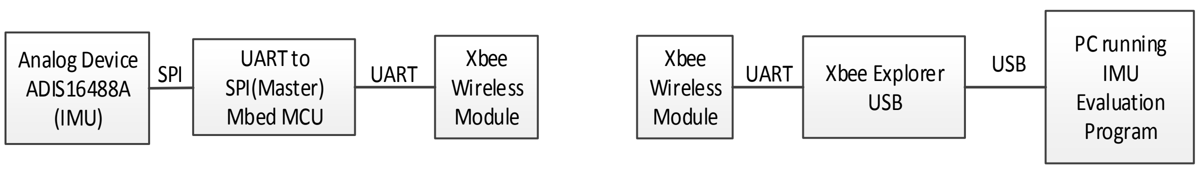


Figure 2. IMU-to-evaluation system wireless communication block diagram

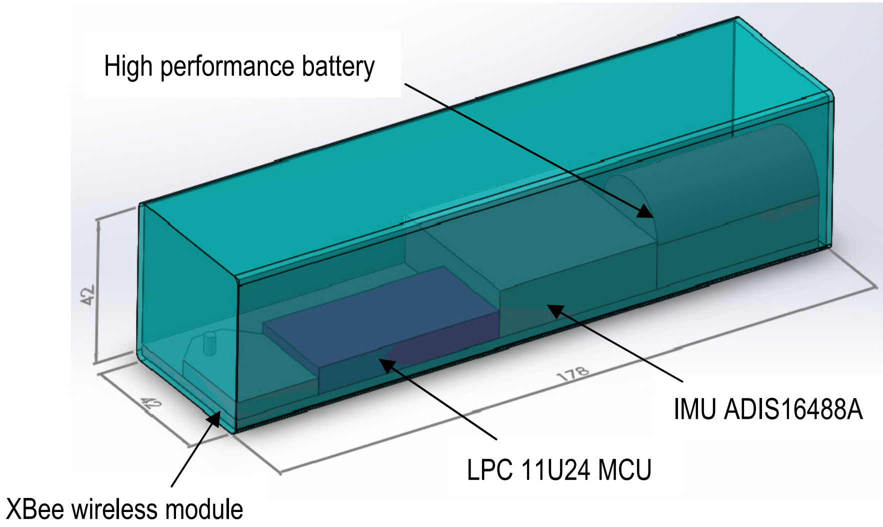


Figure 3. Design of wireless MEMS-based IMU navigation module. All dimensions are in mm

B. Proposed downhole wireless transmission

ZigBee protocol station-based data transmission is proposed for downhole wireless communication. ZigBee was introduced in 1998, standardized in 2003, and revised in 2006, aiming at highly reliable, cost-effective, secure, global wireless data transmission standard and low power consumption [Gislason, 2008]. To meet the low power and low cost criteria, relatively low data transmission rate is a constraint. Nevertheless, it is still very large at 250,000 bits/sec compared to the MPT or EMT with maximum of 100 bits/sec. Not only the modern and simple specifications of IEEE protocol 802.15.4 make it reliable, but the mesh networking feature further enhances its reliability. With mesh networking, data from the first node is able to reach any other node within allowable distance in the ZigBee network. The prices on the market for low-volume sales are around a few dollars, which is low enough for the oil industry. Moreover, as the scope of the 802.15.4 market grows, some have predicted that this radio market will hit the bottom (below $ 1 in quantity) in the next 3-5 years [Gislason, 2008]. ZigBee can operate for years on a pair of AA batteries due to its low power consumption. One of the reasons for this low power feature is that radios and microcontrollers can sleep, since a node on the ZigBee network does not require a constant contact with the network to remain part of it. ZigBee-based wireless modules are disposed in the hollow interior of a drill pipe, separated by some desired distances. Downhole data is transmitted by an initial transmitter at the BHA, relayed though a numerous closed autonomous wireless modules, all powered by local batteries, and eventually received by the receiver on the surface [Varveropoulos & Taherian, 2011].

Another key benefit of such system is that it allows the entire drill string to act as a local network within the BHA, and many downhole tools located somewhere inside the drill string can be individually addressed and/or turned on and off.

C. Wireless navigation module prototype

The discussed wireless navigation module design has been implemented and tested on a solderless breadboard as shown on Figure 4a, from left to right: MBED LPC11u24 MCU, AD ADIS16488A IMU, and XBee Pro S1 RF transmitter. The module was powered using a commercial 9 V battery. The receiver (Figure 4b) connected with a computer continuously and received data from the wireless IMU module at a sampling rate of 5 Hz, which could be considered adequate for the very slow drilling processes and displayed and logged the accelerometers and gyroscopes measurements through the terminal of the Tera Term software. The wireless IMU data transmission was very stable and achieved excellent data integrity when tested within a 30-meter drill pipe.

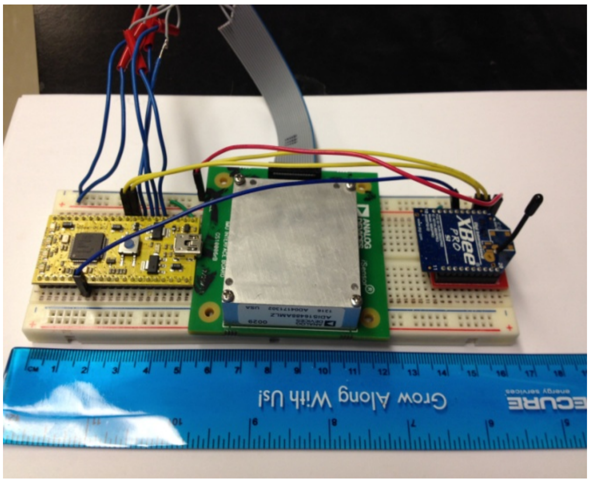


Figure 4a. Wireless navigation module prototype

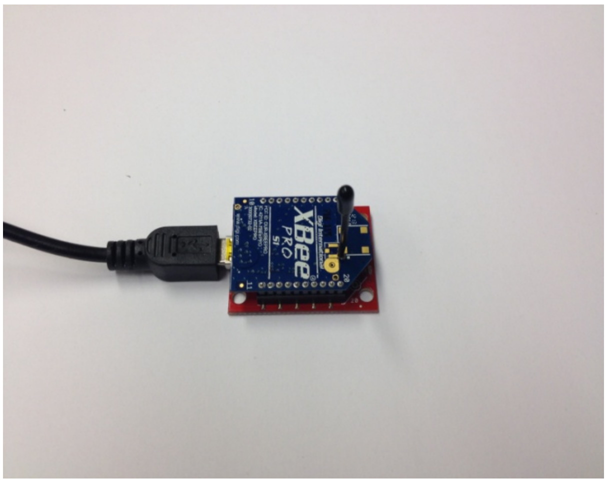


Figure 4b. XBee Pro S1 receiver with XBee Explorer USB adapter

Conclusion

The proposed wireless data transmission for a downhole IMU-based R-IDA setup makes precise controlled rotation during R-IDA error compensation procedures possible, since rotating an IMU with wired cables would be impractical and would result in constant forces and torques applied to the wires, leading to uncontrollable noise induction. This approach makes the practical implementation of the IDA method downhole feasible, which can lead to the replacement of traditional magnetometer or compass-based navigation with modern inertial navigation systems.

Acknowledgement

This study was supported in part by the Natural Sciences and Engineering Research Council of Canada and PetroJet Canada Inc.

Bibliography

[Analog Devices, 2014] Analog Devices. "ADIS16488A: Ten Degrees of Freedom MEMS Inertial Sensor with Dynamic Orientation Outputs." Internet: http://www.analog.com/en/mems-sensors/mems-inertial-measurement-units/adis16488a/products/product.html, [Mar. 30, 2014]

[Baker Hughes, 2014] Baker Hughes, "E-MTrak Electromagnetic Telemetry." Internet: http://www.bakerhughes.com/products-and-services/drilling/drilling-services/measurement-while-drilling/e-mtrak-electromagnetic-telemetry, [May 15, 2014].

[Bosworth et al, 1998] S. Bosworth, H. S. El-Sayed, G. Ismail, H. Ohmer, M. Stracke, C. West, and A. Retnanto. "Key issues in multilateral technology." Oilfield Review, no. 4, pp. 14-28, 1998.

[Farahani, 2011] S. Farahani, ZigBee Wireless Networks and Transceivers. Newnes: Oxford, UK, 2011.

[Gao et al, 2006] L. Gao, D. Finley, W. Gardener, C. Robbins, E. Linyaev, and J. Moore. "Acoustic telemetry delivers more real-time downhole data in underbalanced drilling operations," in Proc. SPE/IADC Drilling Conf., Feb. 2006.

[Gislason, 2008] D. Gislason, ZigBee Wireless Networking. Newnes: Oxford, UK, 2008.

[Jellison et al, 2003] M. Jellison, D. R. Hall, D. C. Howard, H. T. Hall, R. C. Long, R. B. Chandler, and D. S. Pixton. "Telemetry drill pipe: Enabling technology for the downhole internet," in Proc. SPE/IADC Drilling Conf., 2003.

[Jurkov et al, 2011] A. S. Jurkov, J. Cloutier, E. Pecht, and M. P. Mintchev, “Experimental feasibility of the in drilling alignment method for inertial navigation in measurement-while-drilling,” IEEE Trans. Instrum. Meas., vol. 60, no. 3, pp. 1080–1090, Mar. 2011.

[Kyle et al, 2013] D. Kyle, M. Fripp, and K. Fink. "Acoustic telemetry for oilfield operations," IEEE Mag. Instrum. Meas., no. 6 pp. 18-24, Dec. 2013.

[National Oilwell Varco, 2014] National Oilwell Varco, “IntelliServ.” Internet: http://www.nov.com/intelliserv/, [May 20, 2014].

[Pasicznyk, 2001] A. Pasicznyk. "Evolution toward simpler, less risky multilateral wells." In Proc. SPE/IADC drilling conf., Mar. 2001.

[PetroWiki, 2013] PetroWiki, “Measurement while drilling (MWD).” Internet: http://petrowiki.org/Measurement_while_drilling_%28MWD%29, [Sept. 12, 2013].

[Shelkholeslami et al, 1991] B. A. Shelkholeslami, B. W. Schlottman, F. A. Siedel, and D. M. Button, “Drilling and production aspects of horizontal wells in the Austin Chalk,” J. Petroleum Technol., vol. 43, no. 7, pp. 773–779, Jul. 1991.

[Torkildsen et al, 2004] T. Torkildsen, I. Edvardsen, A. Fjogstad, A. Saasen, and T. Omland, “Drilling fluid affects MWD magnetic azimuth and wellbore position,” in Proc. SPE/IADC Drilling Conf., Dallas, TX, pp. 595–602, 2004.

[Varveropoulos & Taherian, 2011] V. Varveropoulos, R. Taherian, “Wireless telemetry through drill pipe”, US Patent 20110018734, Jan. 27, 2011

[Wang et al, 2013] Z. H. Wang, M. Poscente, D. Filip, M. Dimanchev and M. P. Mintchev, “Rotary in-drilling alignment using an autonomous MEMS-based inertial measurement unit for measurement-while-drilling processes,” IEEE Mag. Instrum. Meas., no. 6 pp. 26-34, Dec. 2013.

[Wasserman et al, 2008] I. Wasserman, D. Hahn, D. H. Nguyen, H. Reckmann, and J. Macpherson. "Mud-pulse telemetry sees step-change improvement with oscillating shear valves," Oil & gas Journal, vol. 106, no. 24 pp. 39-40, 2008.

Authors' Information

Zhenhua (Robin) Wang *is with the Department of Electrical and Computer Engineering, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; e-mail: zhewang@ucalgary.ca. Major Fields of Scientific Research: Oilfield drilling navigation and instrumentation.*

Tao Li *is with the Department of Electrical and Computer Engineering, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada. Major Fields of Scientific Research: Downhole navigation and instrumentation.*

Myles McDougall *is the founder and President of PetroJet Canada Inc., Calgary, AB, Canada, website: http://www.petrojet.ca/leadership-team.php. He obtained a BA degree in Economics from the University of Calgary and a joint MBA from the University of Western Ontario/ IESE - University of Navarra, Barcelona, Spain.*

Dan McCormack *is the Vice President of PetroJet Canada Inc., Calgary, AB, Canada, website: http://www.petrojet.ca/leadership-team.php. He obtained a B. Sc. degree in Mechanical Engineering from the University of Alberta.*

Martin P. Mintchev *is with the Department of Electrical and Computer Engineering, and the Faculty of Medicine, University of Calgary, 2500 University Dr. NW, Calgary, AB, T2N 1N4, Canada; e-mail: mintchev@ucalgary.ca. Major Fields of Scientific Research: Biomedical engineering, Oilfield instrumentation.*

INFORMATION TECHNOLOGY OF PROCESSING INFORMATION OF THE CUSTOMS CONTROL

Borys Moroz, Sergii Konovalenko

Abstract: *The paper considers the issues and highlighted the development of information technology processing of information of customs control, which uses effective methods and tools to identify risks of violation of customs legislation. It was suggested take a holistic mechanism for the preparation, processing and interpretation of the results. The proposed information technology will increase the efficiency of the automated system of analysis and risk management Customs Service of Ukraine.*

Keywords: *customs control, risk management, neural networks*

ACM Classification Keywords: *I.5 Pattern Recognition – I.5.1 Models – Neural nets*

Introduction

Efficient operation of the Customs Service of Ukraine is possible only if the use of modern information technologies, methods and tools which will allow faster and better to solve the problem posed by the State [Пашко, 2008]. As is not uncommon to provide economic agents false information about the characteristics of goods for understatement (overstatement) of the customs value, transportation of smuggling, the use of high-quality business intelligence will improve the fight violation of customs legislation [WCO]. The introduction of such technologies is carried out in accordance with international standards and with the maximum involvement of the scientific and technical potential of the Customs Service of Ukraine, which allows you to select a part of the current research is the development and application of methods and tools for processing of information customs control.

Currently there are many methods and tools are used to develop decision support systems, classification, recognition, but from this mass want to highlight methods of artificial intelligence, model that use the neural networks. The theory of neural networks in the last decade gained a lot of practical applications in various fields of science and technology (face recognition, image, medicine and diagnostics, analysis stock and currency markets, using machines and robots). Such a wide distribution and availability of the many software packages modeling gives us reason to assume the possibility of using this mathematical model for the needs of the Customs Service of Ukraine.

Publications on the application of information processing methods of customs control and relatively few, mostly, they are conceptual in nature, highlighting the issues and relevance of the topic [Пашко, 2011]. In [Семенко, 2008] describes the possibility of applying the theory of fuzzy sets for the purpose of risk analysis, which allowed using membership functions to operate more flexibly with the calculation of risk. In our case, to ensure the quality of business intelligence is required to provide an opportunity to study and adapt the system that directs our choice in favor of neural networks.

Problem definition

The purpose of the article is to consider the theoretical and practical aspects of the development of information technology information processing customs control for risk analysis violations of customs legislation, in connection with which there is a need to solve the following problems:

1. To consider component parts of information technology;
2. To describe selected methods and means processing information of customs control;
3. To analyze effectiveness of the proposed system as a whole;
4. Based on the analysis to advice on possible application of the developed model.

Customs Risk Management – Economic security Ukraine

Development of world trade and foreign economic activity dictate their conditions to facilitate customs procedures during the import and export of goods, reducing of risks violation of economic security. In this regard, the implementation of customs emphasis shifted in favor of the application of risk management, the use of new technologies of nondestructive testing, close cooperation with customs services of other countries. This enables optimal use of available resources and the Customs Service to make the control of customs clearance in those areas where there is the greatest risk of violations, allowing the bulk of goods and individuals relatively free passage of customs control.

Consider in detail the system of risk analysis and management. Risk analysis violations of customs legislation is implemented using risk profiles (RP). RP applied during customs control and customs clearance of goods and vehicles crossing the customs border, and have a purpose to warn inspectors of a possible risk violations of the law during a specific foreign trade operations. Developing RP includes implementation of such actions:

- Definition of risk indicators;
- Selectivity rating;

Evaluation of the significance of this RP to fill budget;

Determine the effect of negative RP stories, etc.

The key concept in the customs risk management is a risk indicator. Indicator of risk - is a definite criterion that is used to identify potential violations of customs laws. This system, using the input data must carry out risk assessments on declaration or situation during customs clearance. If there is a risk, the system makes recommendations on the application of the necessary measures to minimize the damage from potential violation of customs legislation. Using a system of risk analysis methods and effective means of intelligent information processing will improve the quality of risk identification customs violations, smuggling.

System of intelligent information processing (SIIP)

Create an information technology processing information of customs control as noted above will use artificial intelligence. The process of developing the intellectual system of information processing (SIIP) generally requires the following steps (Figure 1):

- Perception of information from different sources (usually a heterogeneous vector of input data);
- Preparation or normalization of the input vector;
- Information processing;
- Interpetation of results.

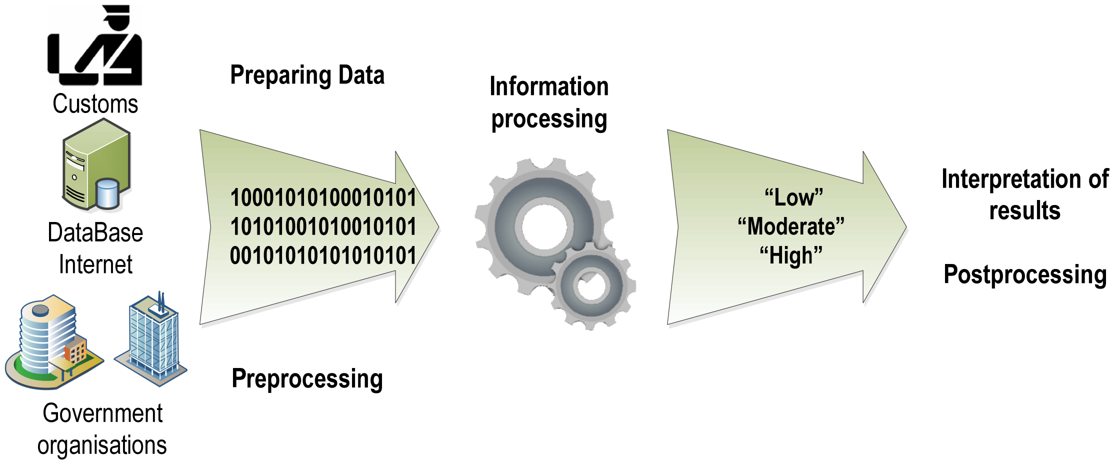


Figure 1. Model of the SIIP

Domain analysis of any problem means processing sufficiently large data sets, wearing a mainly polytypic nature and allocation factors and traits of the most important [Byuyul, 2005]. Since in most cases the information processing system receives input data for analysis from various sources, it becomes necessary to bring them to a suitable format for consideration. The primary source of data may make the storage and database of commercial and government organizations submitted documents, the internet network, i.e. much information as possible, which can be useful for decision making. Given the fact that intelligent systems have the properties of learning, it is important to pay attention to pretreatment and preparation of input data sets. Without doing this, we deteriorate the quality of information analysis system (pattern recognition, classification, etc.), and in some cases it will even fail to perceive the input vector data. The process of intelligent information processing depends on the choice of methods, which we believe meet the requirements to achieve the objective. Perhaps above all it is the methods using machine learning algorithms.

Finally, the interpretation of the results puts a goal before the reports in a clear formula for the user of the result of recognition or classification. In the context of risk analysis, violation of the customs legislation, this level of risk - "High", "Moderate" or "Low".

Preparing input vector "Information of customs control"

Preprocessing is the procedure for preparing data for analysis, during which they are in compliance with the requirements determined by the specifics of the problem being solved (subject domain) and used model processing (analysis) of the information received. Typically, data preprocessing includes two components [BaseGroup]:

1. Cleaning and optimization;
2. Transformation, normalization.

Cleaning is done in order to eliminate factors that reduce the quality of data and analytical algorithms hindering work. It includes processing duplicates, contradictions and fictitious values, restoration and filling gaps, smoothing and cleaning data from the noise suppression and editing of the abnormal value. In addition, the cleaning process restores the structure, completeness and integrity of the data, converted incorrect formats.

Optimization of data as part of pre-treatment includes reduction the dimension of the input data, the identification and elimination of non-significant features. The main difference between the optimization of purification, that the factors that are fixed in the cleaning process, significantly reduce the accuracy of the solution, or do the work of analytic algorithms impossible. Problems to be solved by the optimization, data adapted to a specific problem and improve the efficiency of their analysis.

With regard to transformation and normalization of data, this step is required to make the information comprehensible to the terms used by the analytical model. This includes operations such as type conversion, quantization, encoding and so on. Each method of analysis requires that the original data were in any particular form. For example, neural networks work only with numeric data, and they should be normalized [Haykin, 1998]. Training data set SIIP must satisfy several criteria [BaseGroup]:

1. The representativeness - data should illustrate the true state of things in the subject area;
2. The consistency - conflicting data in the training set will result in a poor quality of the network training.

Table 1. Forming the input vector "Information of customs control"

№	Identification characteristics	Data types	Accepted values	Coding value
X_0	Country of Origin	string	Offshore	00 (bin)
			The EU countries	01 (bin)
			The EEA countries	10 (bin)
			other countries	11 (bin)
X_1	Product Code	integer	In accordance with the classifier	Range [0…1]
X_2	Customs cost	float	In accordance with the customs declaration	Range [0…1]
X_3	Quantity of goods	integer	Number of units or batches of delivery	Range [0…1]
X_4	Weight of goods	float	Weight unit of goods or the supply of the party	Range [0…1]
X_5	Invoice cost of a product	float	In accordance with the customs declaration	Range [0…1]
X_6	The difference gross and net product	float	no more 5%	0
			from 5% to 8%	0.5
			more than 8%	1
X_7	The history of the participant of foreign economic activity	string	black list	0
			gray list	0.5
			white list	1

These criteria provide for himself the whole complex of actions preprocessing of different types of input data. As a rule, neural networks like multilayer perceptron using sigmoid activation function. That is, the input vector $x = [x_0, x_1, \ldots, x_i]^T$ must be brought to the range [0...1] or [-1...1] by (1) or (2).

$$x_i = \frac{x_i}{\max(X)}. \quad (1)$$

$$x_i = \frac{x_i - \mu}{S}, \quad (2)$$

where S – range ($\max(X) - \min(X)$ or standard deviation, μ – average value.

If you need to partition the continuous value into segments of equal length, the quantization can be performed as the initial value of the division by a constant value (quantization step) and the integral part of the quotient:

$$y_q = \frac{y - y_0}{h}, \quad (3)$$

where h – quantization step.

An example of the process of forming the input vector is presented in the Table 1.

The components of the input vector are encoded by the following principle:

1. The variable X_0 is encoded binary numbers;
2. The variables $X_1 \ldots X_5$ are transformed by the formula (1) to the range [0 ... 1];
3. Variables $X_6 \ldots X_7$ take only three values, so we put them in compliance with three numeric values of {0, 0.5, 1}. These values match the level of risk {"Low," "Moderate," "High"} [ASYCUDA], [Konovalenko, 2012].

In such a way input vector was transformed to a common format and range – [0 ... 1], which is suitable for the activation function. Now it can be input into the neural network used for training and classification.

Information processing in customs control

System of intelligent information processing of customs control uses the methods and means of artificial intelligence. It was suggested [Moroz, 2011] the mathematical apparatus of artificial neural networks to detect the risk of violating the customs legislation. Quality training SIIP is probably the most important task, it is necessary to solve the developer, because it directly influences on the result [Bishop, 2007]. So first of all, you need the original set of input data separated on:

1. The training set $\left\{\left(x^{(1)}, y^{(1)}\right) \ldots \left(x^{(m)}, y^{(m)}\right)\right\}$;
2. Cross validation - $\left\{\left(x_{cv}^{(1)}, y_{cv}^{(1)}\right) \ldots \left(x_{cv}^{(m_{cv})}, y_{cv}^{(m_{cv})}\right)\right\}$;
3. Test - $\left\{\left(x_{test}^{(1)}, y_{test}^{(1)}\right) \ldots \left(x_{test}^{(m_{test})}, y_{test}^{(m_{test})}\right)\right\}$.

The distribution needs to be done in such proportions - respectively 60%, 20% to 20% of that total amount. This will allow us to use the $J_{train}(\Theta)$ (*Training error*), $J_{cv}(\Theta)$ (*Cross Validation error*) and $J_{test}(\Theta)$ (*Test error*). The general form of the error function (target function) has the following form:

$$J(\Theta) = \frac{1}{2m}\left[\sum_{i=1}^{m}\left(h_\Theta\left(x^{(i)}\right) - y^{(i)}\right)^2 + \lambda\sum_{j=2}^{n}\Theta_j^2\right], \quad (4)$$

where $h_\Theta(x) = g(\Theta^T x)$ (model that can be configured, or hypothesis), Θ - vector of parameters, which is configured, λ - parameter of regularization member.

Perceptron learning process gradient methods provides an iterative algorithm to minimize the target function $J(\Theta)$ by adjusting the parameters Θ (5).

$$\Theta_j = \Theta_j - \alpha\left(\frac{1}{m}\sum_{i=1}^{m}\left(h_\Theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)} + \frac{\lambda}{m}\Theta_j\right), \quad (5)$$

where α - learning rate.

Learning rate α () must be chosen so as to avoid these problems [Ng, 2012]:

- If α too small, the gradient methods are quite slow to converge;
- If α too large, then it is possible that optimization process reusable "jump over" minimum of the target function, or even completely diverge.

During optimization, it is advisable to plot the target function $J(\Theta)$ of the number of iterations and stop the learning process when the decrease in the values of a single iteration will not be less than a certain threshold, for example 10^{-3}. Ideally, it depicts the relaxation process shown in Figure 2. b) During training SIIP appropriate and probably necessary to calculate the above errors ($J_{train}(\Theta)$, $J_{cv}(\Theta)$ and $J_{test}(\Theta)$) in the space of the corresponding sets. Analyzing these errors can be identified several issues that affect the ability of the classifier to generalize. These problems can be described by the following terms:

1) *Under fit* - λ has big value (Figure 2. a));
2) *Over fit* - $\lambda = 0$ (Figure 2. a));
3) *Just right* - λ has average value (Figure 2. a)).

The dependencies of the error function of the regularization parameter λ and capacity set *m* (Figure 3) reveals the insidious problem of over fit, i.e. situation where the classifier works correctly only on the training set.

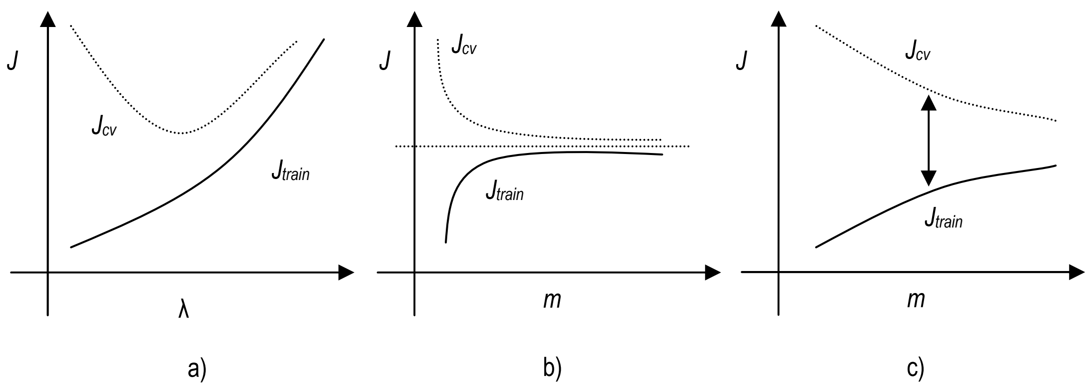


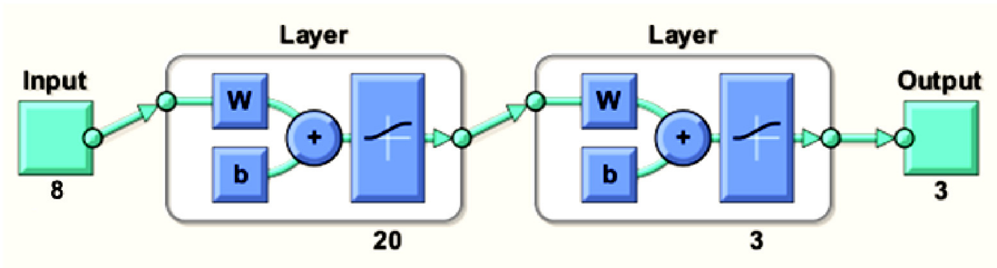
Figure 2. a) Dependence $J(\lambda)$, b) and c) Dependence $J(m)$, where *m* – capacity of set

If during training and cross-validation we get a big mistake (Figure. 2. b), increasing the number *m* not solve the problem. In the situation (Figure 2. b)), the opposite - increase slightly improve the quality of education in general [Ng, 2012]. In addition, this procedure can add dimension reduction of the input space (if needed).

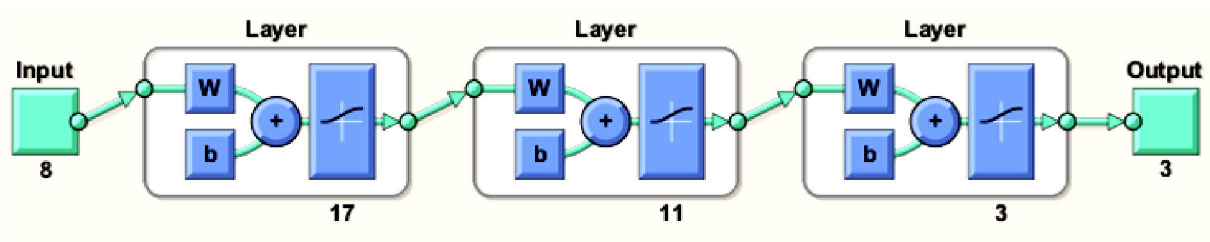
After spending 6 experiments for various architectures of multilayer perceptron, revealed that quasi-Newton method (BFGS) minimization of target function yields the Levenberg-Marquardt (LM) for neural network training time and quality of recognition in the test set. Based on the results of experiments allocated optimal neural network model for one or two hidden layers:

1. Architecture with one hidden layer - 8-[20]-3, learning method LM, recognition error 1.4%;

Architecture with two hidden layers - 8-[17-11]-3, learning method LM, recognition error 1.2% (Figure 3 b)).



a)



b)

Figure 3. Neural model of the SIIP

So we got a complete description of effective models processing of information of customs control, which allow identifying the risk violation of customs legislation. And as a result will improve the quality customs clearance of goods, while reducing costs at the time.

Conclusion

As a result of this work was to develop a holistic model of risk identification violations of customs legislation on the basis of the type of neural network multilayer perceptron.

The paper discusses methods of converting continuous and discrete types of data suitable for analysis of the vector and sets. On the example domain "Information customs control" was shown how training set for the recognition of risk violations of customs legislation.

Were identified analysis techniques quality of the learning process SIIP and possible ways to overcome such problems of neural networks, as loss of the ability to generalize. Taking into account these recommendations improves learning algorithm and thus affect the quality of recognition risks violation of customs legislation.

On the basis of this experiments and comparative analysis of the quality of recognition of input images, highlighted the optimal architecture of the neural network.

Further research should be devoted to the study of effective methods of calculating the optimum multilayer perceptron architecture, as well as the mechanism of improvement opportunities neuroclassifier by sharing with fuzzy logic, using genetic algorithms.

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

[ASYCUDA] Automated SYstem for CUstoms DAta (ASYCUDA). Risk Management [Online]. Available: http://www.asycuda.org

[BaseGroup] BaseGroup Labs. Technology of data analysis [Online]. Available: http://www.basegroup.ru

[Bishop, 2007] Christopher M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). - Springer, 2007. - P. 738.

[Byuyul, 2005] A. Byuyul, P. Tsefel, SPSS: Art of information processing, (in Russian), DiaSoft, 2005.

[Haykin, 1998] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed., Prentice Hall, 1998.

[Konovalenko, 2012] Konovalenko Sergii. Preprocessing "raw" data sets as an important aspect of intelligent information processing, International Book Series "Information Science and Computing". Artificial Intelligence Driven Solutions to Business and Engineering Problems. - Rzeszow and Sofia : ITHEA, 2012. - Book 27. - P. 133-139.

[Moroz, 2011] B. Moroz, S. Konovalenko, "Aplication of neural networks within of the concept E-customs," (in Russian), in Information Science and Computing, book 22, Applicable Information Models, K. Markov, V. Velychko, Ed., 1st ed. Sofia, Bulgaria: ITHEA, 2011, pp. 104-110.

[Ng, 2012] Andrew Ng Machine Learning Online Course [Online]. Available: http://www.coursera.org.

[WCO] World Customs Organization. WCO course on risk assessment, profiling and targeting [Online]. Available: http://www.wcoomd.org

[Пашко, 2008] Пашко П.В. Основи митної справи в Україні: Підручник – К.: Знання, 2008. – 652 с.

[Пашко, 2011] Митні інформаційні технології: навч. посібник / О. Ф. Волик, О. В. Кащеєва, І. В. Дорда та ін.; За ред.

[Семенко, 2008] Семенко, О. М.-М. Підвищення ефективності роботи по протидії контрабанді та порушенням митних правил, митного контролю за рахунок впровадження автоматизованої системи аналізу та управління митними ризиками на основі Fuzzy - технології: до вивчення дисципліни / О. М.-М. Семенко. – К.: LAT & K, 2008. – 238 с.П. В. Пашка. - К. : Знання, 2011. - 391 с.

Authors' Information

***Borys Moroz** – Professor, doctor of technical sciences, counselor of customs service of 1st rank, dean of faculty information and transport systems and technologies of the Ukrainian Academy of Customs, 2/4 Dzerzhinsky str., Dnipropetrovs'k, 49044, UKRAINE; e-mail: realtime07@mail.ru*

Major Fields of Scientific Research: information theory, fuzzy logic

***Sergii Konovalenko** – inspector of customs service of 3rd rank, head of the laboratory of information systems and processes in customs affair of the Ukrainian Academy of Customs, 2/4 Dzerzhinsky str., Dnipropetrovs'k, 49044, UKRAINE; e-mail: customslab@rambler.ru*

Major Fields of Scientific Research: Computational neuroscience, computational bioinformatics

GEOMETRIC APPROACH FOR GAUSSIAN-KERNEL BOLSTERED ERROR ESTIMATION FOR LINEAR CLASSIFICATION IN COMPUTATIONAL BIOLOGY

Arsen Arakelyan, Lilit Nerisyan, Aram Gevorgyan, Anna Boyajyan

Abstract: *Classification and feature selection techniques are among the most commonly used mathematical approaches for analysis and interpretation of biological data. One of the important characteristics of any classifier is its classification error, which is important to take into consideration for accurate data analysis. The most popular error estimation techniques (resubstitution, bootstrapping, cross-validation) strikingly vary in performance. It is well known that more accurate classifiers such as bootstrapping, cross-validation are very slow, while heavily biased resubstitution is very fast. Recently, a new bolstered error estimation technique has been proposed that optimally combines speed and accuracy. It uses a Monte-Carlo sampling based algorithm for classification for the general case, but for the case of linear classification, an analytical solution may be applied. In this paper we introduce geometric approach for bolstered error estimation and compare its performance with other error estimation algorithms. The results obtained show that geometric bolstered error estimation algorithms are very fast error estimation techniques characterized by accuracy comparable with LOO and having lower variance. These algorithms are useful for analyzing extremely large numbers of features and may find their applications in wide fields of - omics data analysis.*

Keywords: *Biology and genetics, Classifier design and evaluation, Machine learning.*

ACM Classification Keywords: *A.0 General Literature - Conference proceedings, I.5.2 – Classifier design and evaluation, Feature evaluation and selection. J.3 - Biology and genetics*

Introduction

Classification techniques have found their wide application in various fields of biomedical research [Sayes, 2007]. The classification problem may be stated as follows: given a set of objects belonging to two or more classes and described by a set of features, the aim is to design a classifier that will correctly predict class memberships of new objects. Support vector machines (SVMs) are among the most popular classifiers widely used for classification and feature extraction in computational biology research. SVMs are used for prediction of protein secondary structure [Nguyen, 2011], analysis of protein and DNA sequences [Choi, 2011; Lee, 2011], protein classification [Cai, 2003], prediction of protein-protein interactions [Cui, 2012], identification of transcription binding sites [Holloway, 2007], and analysis of gene expression data [Golub, 1999; Maulik, 2013].

Occasionally, a classifier may fail to correctly assign the membership of new objects resulting in classification error. Classification true error is the error rate of the classifier if it was tested on the true distribution of cases [Nolan, 1997]. Since the true distribution is generally unknown, there is a need to come up with a proper estimation of the classification true error. Considering its importance both for assessment of the classifier itself and for accurate interpretation of classification results, several algorithms have been developed for classification true error estimation.

The best algorithm for approximation of the true error is considered to be the hold-out estimation [Nolan, 1997], where the dataset is divided into independent training and test sets. The purpose of the training set is to design the classifier, while the test set is used for assessing the classification error. This method, however, requires large datasets, which are not always available, especially in the field of genomics. In high-throughput gene expression analysis, researchers often deal with very small sample sizes, making the application of the above mentioned strategy basically impossible [Allison, 2006]. In such cases, training set based error estimation approaches, such as resubstitution (Resub), leave-one-out cross-validation (LOO) and bootstrapping methods (BST), are commonly used. These techniques are shown to be strikingly different in terms of speed and accuracy [Dougherty, 2010].

Resub uses the whole training data to estimate the error of a classifier, and is considered to be the fastest among available algorithms [Devroye, 1996]. However, it has been shown that it is heavily biased, especially in the case of small sample sized settings [Devroye, 1996].

LOO error estimation is a case of cross-validation algorithms when a single observation from the original sample is used as validation data, and the remaining observations – as training data. This is repeated until each observation in the sample is used as validation data. LOO error estimation is shown to be nearly unbiased, but to have large variance. Moreover, the speed of LOO slows with the increase of sample size [Lachenbrucha, 1968].

BST error estimation is based on generation of a test set from the training set using sampling with replacement technique [Efron, 1983]. For correct error estimation by bootstrapping, it is suggested to use 100-200 bootstrap samplings [Efron, 1983]. Bootstrapping error estimation is usually pessimistically biased, but has lower variance compared to LOO. In addition, bootstrapping is very slow compared to Resub and LOO.

Recently, Braga-Neto and Dougherty [Braga-Neto, 2004] have proposed another error estimation technique called “bolstered error estimation” (BOL). The principal idea of this approach is the following. The available data points are used as a training set for the design of the classifier. Next, a bolstered distribution of these points is generated based on class-dependent variances. In simple cases, bolstering is performed by constructing p-dimensional spheres around the points (spherical bolstering), where p is the number of features (Figure 1).

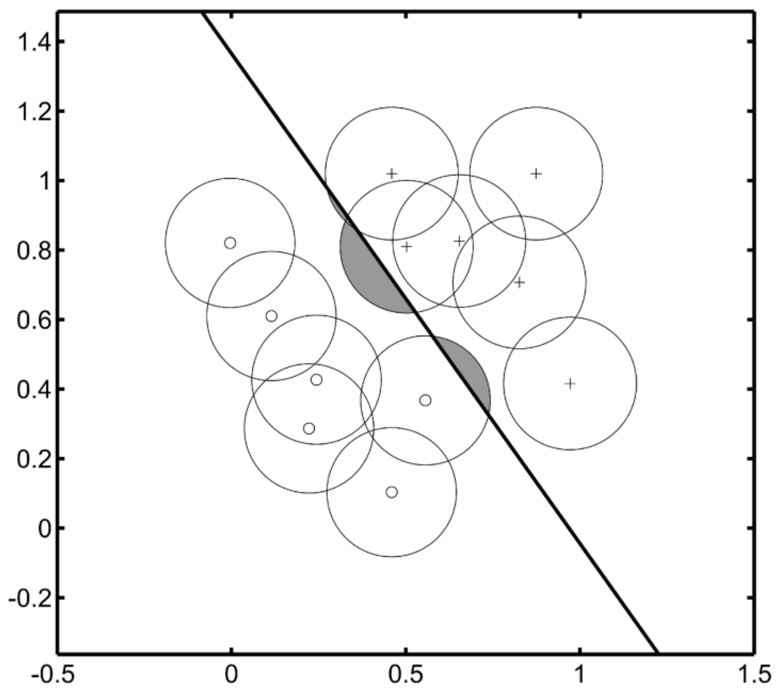


Figure 1. Graphical representation of bolstered error estimation. Shaded areas of the circles represent error contributions of the points. The average of all the error contributions is the bolstered error estimate

The test set is generated by taking new points from these spheres. The use of bolstered space increases the accuracy of error estimation. According to the authors [Braga-Neto, 2004], bolstered error estimation combines high computational speed of resubstitution and the accuracy of LOO algorithms. In addition, they have proposed a semi-bolstering technique, when bolstering is applied only on correctly classified points. In the original paper [Braga-Neto, 2004], bolstering error estimation is calculated using Monte-Carlo integration (mBOL), however, in the case of linear classification and spherical bolstering, it is possible to find analytical solutions for error estimation that may be more accurate and less computationally intensive.

In this paper we introduce an analytical approach for bolstered error estimation for linear classification with SVMs based on computational geometry approaches.

Methods

Geometric bolstered error estimation algorithm (gBOL)

The algorithm is designed for linear classification and spherical bolstering. It proceeds by firstly training the linear classifier (e.g. SVM) based on dataset points. Next, spherical bolsters are generated around the points, with sphere radiuses being equal to the variance in each class, and sphere dimensions representing the number of features (p). If the signed distance of a sphere center from the hyperplane is less than the sphere radius, the hyperplane may cut the sphere resulting in generation of a spherical cap. The ratio of the cap volume to the whole volume of the sphere is the error contribution of the given point. In the most extreme cases the whole sphere may appear in the opposite decision region, resulting in ratio equal to 1 (100% misclassification). Finally, classification error is estimated by computing the ratio of the volume of spherical caps or spheres appearing in the opposite decision region to the overall volume of the spheres (Figure 2).

The pseudocode for gsBOL calculation:

Let N data points $\{x_i, y_i\}$ for $i = 1...N, x \in R^p$ and classes $y = \{-1,1\}$ be the training set.

Set error *err* = *0*;

Train the linear classifier;

For each class in y

  Calculate the radius **σ** of the spherical bolstering kernel;

  Calculate the volume of the sphere V with radius **σ**;

    For each point belonging to the class in y

      Calculate the signed distance d_i from the point x_i to the separating hyperplane;

      Calculate the volume of the spherical cap C_i in opposite decision region using **σ** and d_i;

      Calculate the bolstered error $errp_i$ for the point x_i as C_i/V;

      Accumulate $errp_i$ in $err = err + errp_i$;

    End

End

The total bolstered error is calculated as *err/N*

Figure 2. The pseudocode for geometric bolstered error estimation

Geometric semi-bolstered error estimation (gsBOL)

The semi-bolstered error estimation algorithm is based on the gBOL algorithm. The difference is that bolstering kernel is calculated only for correctly classified objects [Braga-Neto, 2004], while misclassified point are assigned 100% error.

Formulas used in calculations

Linear support vector machines

For N data points $\{x_i, y_i\}$ for $i = 1 \ldots N, x \in R^p$ and classes $y = \{-1, 1\}$, the linear support vector solution should satisfy the following conditions [Vapnik, 1995]:

$$\begin{cases} W(x_i) = w^T x_i + b \geq +1, \text{if } y_i = +1 \\ W(x_i) = w^T x_i + b \leq -1, \text{if } y_i = -1 \end{cases} \quad (1)$$

where $w^T x_i + b = 0$ is the separation hyperplane. By minimizing $w^T w$, the margin between both classes is maximized. The signed distance $\boldsymbol{d_i}$ from a data point x_i to the separating hyperplane equals $d_i = W(x_i)\|w\|$.

Bolstering kernel radius

Bolstering kernel radiuses were calculated according to the approach given in [Braga-Neto, 2004]. The Gaussian *p*-variate (*p* is the number of features) zero-mean bolstering kernel generates spheres centered at the data points. The radius of the sphere, which is equal to the standard deviation σ for a given class, can be easily calculated from Euclidean pairwise distances between data points in the given class multiplied by a correction factor. The correction factor is the inverse of χ^2 cumulative distribution at point 0.5 with degrees of freedom equal to *p* [Braga-Neto, 2004].

Cap volume computation and bolstered error estimation

If the distance from a given data point to the separation hyperplane is less than σ, then the hyperplane will divide the bolstering sphere into caps, one of which will be located in the opposite decision region. The volume of this spherical cap actually represents the error contribution of a given data point.

Spherical cap volume can be computed using formulas introduced by Li [Li, 2011] as follows:

$$\begin{aligned} V_n^{cap}(r) &= \int_{\theta-\varphi}^{\theta-0} V_{n-1}(r \sin\theta) d \cos\theta = \\ &= 1/2 \, \mathrm{V}_n(r) I_{\sin^2\phi}\left(\frac{n+1}{2}, \frac{1}{2}\right) \end{aligned} \quad (2)$$

where n is the sphere dimension, r is the radius of the sphere, $I_{\sin^2\phi}\left(\frac{n+1}{2}, \frac{1}{2}\right)$ is the incomplete beta function and $0 \leq \phi \leq \pi/2$ is the colatitude angle.

Algorithms and scripts for Monte-Carlo sampling (mBOL), LOO and BST error estimation

The algorithms and MATLAB codes for BST, LOO, and mBOL error estimation were obtained from previously published sources [Efron, 1994; Hastie, 2009] and [Okun, 2011].

The scripts for gBOL and gsBOL as well as the general framework for comparison of abovementioned estimation techniques were written in MATLAB. For SVM linear classification, MATLAB's built-in svmtrain function was used.

The Matlab source code for geometric bolstering is located at

http://www.mathworks.com/matlabcentral/fileexchange/40118/.

All the used scripts are available at http://www.molbiol.sci.am/big/jbc/Arakelyan_JBC_src.zip.

Datasets

For evaluation of different error estimation techniques we have used two datasets from previously published results on microarray-based classification of non–small cell lung cancer [Showe, 2009] and prediction of survival in sub-optimally debulked patients with ovarian cancer [Bonome, 2008]. These datasets are available in Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under identifiers GSE13255 and GSE26712, respectively.

In the first study authors have identified a 29-gene signature that separates these two patient classes with 86% accuracy [Showe, 2009]. The classification was performed using an SVM with recursive feature elimination with 10-fold cross-validation. In the second study, expression profiles of 57 genes were used for prediction of survival in sub-optimally debulked patients. Classification was performed with hierarchical clustering and 10-fold cross-validation [Bonome, 2008].

Additionally, we have generated synthetic data (200 samples per class, 5 dimensions) drawn from Gaussian class-conditional distribution with means equal to -0.5 and 0.5 and spherical variance equal to 2, with Bayes error equal to 0.29. This very high error rate was chosen to evaluate the performance of error estimators in very extrim settings.

Experimental setup

We have assessed the performance of gBOL, gsBOL, as well as Resub, LOO and BST error estimation algorithms according to the following setup:

1. For all datasets we have performed classification error estimation using p = 1, 2, 3, and 5 top ranked genes.
2. All calculations were performed for samples of size n = 5, 10, 20 and 50 per class. For the GSE26712 dataset, n = 50 was not used because the maximum sample size among the classes was 45.
3. The true error for each experiment was calculated using hold-out estimation: In each experiment a sample of size n was independently drawn from the pool. This sample was used as a training set, while the remainder was used as a test set.
4. Each experiment was performed 1000 times.
5. The true error deviation distribution parameters (mean, SD and root mean square (RMS)) were calculated for each experiment.

Results

We have performed an evaluation of geometrical bolstered and semi-bolstered error estimation using synthetic and experimental data, and compared the results with other commonly used error estimation algorithms, as well as the original Monte-Carlo based bolstering algorithm.

Simulation data

The accuracy of error estimation techniques for the synthetic data is shown in Table 1. Because the true error of the synthetic data was known we also evaluated the performance of hold-out estimator, which is thought to be the most accurate true error estimator. However, we found that hold-out estimator was not the best in our experiments. Simple average rank calculation of true error differences showed that mBOL, gBOL and BST are top tree error estimators.

Table 1. Performance of the error estimation algorithms on the synthetic data

		Class 1: Mean = -0.5, variance: 2; Class 2: Mean = 0.5, variance: 2 True Error = 0.29						
		Hold out	**Resub**	**LOO**	**BST**	**mBOL**	**gBOL**	**gsBOL**
		M_{diff} ± SD (RMS)						
p = 1	n = 5	-0.16 ± 0.07 (0.17)	-0.07 ± 0.12 (0.14)	-0.17 ± 0.20 (0.26)	-0.03 ± 0.10 (0.10)	-0.08 ± 0.09 (0.12)	-0.07 ± 0.10 (0.12)	-0.12 ± 0.12 (0.17)
p = 1	n = 10	-0.15 ± 0.06 (0.16)	-0.09 ± 0.09 (0.13)	-0.14 ± 0.14 (0.20)	-0.08 ± 0.07 (0.11)	-0.10 ± 0.08 (0.13)	-0.10 ± 0.08 (0.13)	-0.13 ± 0.09 (0.16)
p = 1	n = 20	-0.13 ± 0.04 (0.14)	-0.11 ± 0.07 (0.13)	-0.13 ± 0.10 (0.17)	-0.10 ± 0.06 (0.12)	-0.11 ± 0.06 (0.13)	-0.11 ± 0.07 (0.13)	-0.13 ± 0.07 (0.15)
p = 1	n = 50	-0.12 ± 0.02 (0.12)	-0.12 ± 0.04 (0.13)	-0.12 ± 0.05 (0.13)	-0.12 ± 0.04 (0.12)	-0.12 ± 0.04 (0.13)	-0.12 ± 0.04 (0.13)	-0.13 ± 0.04 (0.13)
p = 2	n = 5	-0.13 ± 0.07 (0.15)	0.02 ± 0.13 (0.13)	-0.15 ± 0.20 (0.25)	0.07 ± 0.09 (0.11)	-0.03 ± 0.09 (0.09)	0.01 ± 0.10 (0.10)	-0.04 ± 0.12 (0.13)
p = 2	n = 10	-0.11 ± 0.06 (0.13)	-0.03 ± 0.10 (0.10)	-0.13 ± 0.15 (0.20)	0.00 ± 0.07 (0.07)	-0.06 ± 0.07 (0.09)	-0.04 ± 0.08 (0.09)	-0.07 ± 0.09 (0.12)
p = 2	n = 20	-0.09 ± 0.03 (0.10)	-0.05 ± 0.07 (0.08)	-0.09 ± 0.09 (0.13)	-0.03 ± 0.06 (0.07)	-0.06 ± 0.06 (0.08)	-0.05 ± 0.06 (0.08)	-0.08 ± 0.07 (0.10)
p = 2	n = 50	-0.07 ± 0.02 (0.07)	-0.05 ± 0.04 (0.07)	-0.07 ± 0.05 (0.08)	-0.05 ± 0.04 (0.06)	-0.06 ± 0.04 (0.07)	-0.06 ± 0.04 (0.07)	-0.07 ± 0.04 (0.09)
p = 3	n = 5	-0.11 ± 0.07 (0.13)	0.09 ± 0.12 (0.15)	-0.12 ± 0.20 (0.23)	0.14 ± 0.08 (0.16)	0.00 ± 0.08 (0.08)	0.07 ± 0.10 (0.12)	0.03 ± 0.12 (0.12)
p = 3	n = 10	-0.08 ± 0.04 (0.10)	0.03 ± 0.09 (0.10)	-0.09 ± 0.14 (0.16)	0.07 ± 0.07 (0.10)	-0.01 ± 0.07 (0.07)	0.03 ± 0.08 (0.08)	-0.01 ± 0.09 (0.09)
p = 3	n = 20	-0.07 ± 0.03 (0.07)	0.00 ± 0.07 (0.07)	-0.07 ± 0.09 (0.11)	0.02 ± 0.06 (0.06)	-0.03 ± 0.05 (0.06)	-0.01 ± 0.06 (0.06)	-0.04 ± 0.07 (0.08)
p = 3	n = 50	-0.05 ± 0.02 (0.06)	-0.03 ± 0.04 (0.05)	-0.06 ± 0.04 (0.07)	-0.02 ± 0.03 (0.04)	-0.04 ± 0.03 (0.05)	-0.03 ± 0.03 (0.05)	-0.06 ± 0.04 (0.07)
p = 5	n = 5	-0.08 ± 0.07 (0.11)	0.19 ± 0.09 (0.21)	-0.10 ± 0.20 (0.23)	0.23 ± 0.05 (0.24)	0.03 ± 0.07 (0.08)	0.16 ± 0.08 (0.18)	0.14 ± 0.10 (0.17)
p = 5	n = 10	-0.04 ± 0.05 (0.07)	0.12 ± 0.09 (0.15)	-0.05 ± 0.14 (0.15)	0.16 ± 0.06 (0.17)	0.02 ± 0.06 (0.06)	0.10 ± 0.07 (0.13)	0.08 ± 0.08 (0.11)
p = 5	n = 20	-0.01 ± 0.03 (0.03)	0.07 ± 0.06 (0.10)	-0.02 ± 0.09 (0.09)	0.10 ± 0.05 (0.11)	0.01 ± 0.04 (0.04)	0.06 ± 0.06 (0.08)	0.04 ± 0.06 (0.07)
p = 5	n = 50	0.02 ± 0.02 (0.03)	0.06 ± 0.04 (0.07)	0.02 ± 0.05 (0.05)	0.06 ± 0.03 (0.07)	0.01 ± 0.03 (0.03)	0.05 ± 0.03 (0.06)	0.02 ± 0.04 (0.04)

Experimental data

For experimental validation, we have used GSE13255 and GSE26712 datasets (see 2.5).

The accuracy of error estimation techniques in all 16 experiments for the GSE13255 dataset (experiment 1) showed that the distributions of the deviation of the error estimates from the true error obtained by gBOL and gsBOL were comparable with the results of LOO, BST and mBOL (Table 2) and outperformed Resub. Meanwhile, gsBOL appeared to be more accurate than gBOL. Moreover, the data obtained showed that gBOL and gsBOL demonstrate much lower variance compared to LOO.

In terms of computational speed, gBOL and gsBOL clearly "beat" mBOL, LOO and BST, being very similar to hold-out and Resub error estimations. While gBOL and gsBOL demonstrated almost no variability in computational speed depending on the sample size. The speed of mBOL, LOO, and, especially, BST dramatically slowed down when the sample size became more than 10 (Figure. 3).

Table 2. Performance of the error estimation algorithms in experiment 1 (GSE13255)

		True errror	Resub	LOO	BST	mBOL	gBOL	gsBOL
Fea-tures	Sample size	M ± SD	M_{diff} ± SD (RMS)					
p = 1	n = 5	0.47±0.05	0.12±0.12 (0.17)	0.04±0.18 (0.18)	0.17±0.09 (0.19)	0.12±0.09 (0.15)	0.12±0.10 (0.16)	0.09±0.11 (0.14)
p = 1	n = 10	0.46±0.03	0.10±0.09 (0.13)	0.05±0.13 (0.14)	0.12±0.07 (0.14)	0.09±0.08 (0.12)	0.09±0.08 (0.12)	0.08±0.09 (0.12)
p = 1	n = 20	0.46±0.02	0.08±0.07 (0.11)	0.06±0.08 (0.10)	0.09±0.06 (0.11)	0.08±0.06 (0.10)	0.08±0.06 (0.10)	0.07±0.07 (0.10)
p = 1	n = 50	0.49±0.03	0.11±0.05 (0.12)	0.10±0.05 (0.11)	0.11±0.05 (0.12)	0.11±0.05 (0.12)	0.11±0.05 (0.12)	0.10±0.05 (0.11)
p = 2	n = 5	0.35±0.08	0.13±0.13 (0.18)	-0.02±0.19 (0.19)	0.17±0.10 (0.20)	0.06±0.11 (0.13)	0.11±0.11 (0.16)	0.07±0.13 (0.15)
p = 2	n = 10	0.33±0.06	0.07±0.10 (0.12)	-0.00±0.13 (0.13)	0.10±0.09 (0.13)	0.03±0.09 (0.09)	0.06±0.09 (0.11)	0.02±0.10 (0.10)
p = 2	n = 20	0.31±0.05	0.04±0.08 (0.09)	0.01±0.09 (0.09)	0.05±0.07 (0.09)	0.02±0.07 (0.08)	0.04±0.08 (0.08)	0.01±0.08 (0.08)
p = 2	n = 50	0.30±0.04	0.02±0.06 (0.07)	0.01±0.07 (0.07)	0.03±0.06 (0.07)	0.02±0.06 (0.06)	0.02±0.06 (0.07)	-0.00±0.06 (0.06)
p = 3	n = 5	0.35±0.08	0.16±0.13 (0.21)	-0.03±0.19 (0.19)	0.22±0.10 (0.24)	0.07±0.10 (0.12)	0.14±0.11 (0.18)	0.11±0.13 (0.16)
p = 3	n = 10	0.32±0.05	0.10±0.10 (0.14)	-0.00±0.14 (0.14)	0.14±0.08 (0.16)	0.04±0.08 (0.09)	0.09±0.09 (0.13)	0.05±0.10 (0.11)
p = 3	n = 20	0.3±0.04	0.06±0.08 (0.10)	0.01±0.09 (0.09)	0.08±0.07 (0.11)	0.02±0.07 (0.07)	0.05±0.07 (0.09)	0.02±0.08 (0.08)
p = 3	n = 50	0.31±0.03	0.06±0.06 (0.08)	0.03±0.06 (0.07)	0.07±0.06 (0.09)	0.03±0.05 (0.06)	0.05±0.06 (0.08)	0.02±0.06 (0.06)
p = 5	n = 5	0.34±0.07	0.24±0.11 (0.27)	-0.04±0.20 (0.20)	0.29±0.08 (0.30)	0.08±0.10 (0.13)	0.21±0.10 (0.23)	0.19±0.11 (0.22)
p = 5	n = 10	0.30±0.05	0.15±0.09 (0.18)	-0.00±0.13 (0.13)	0.19±0.07 (0.21)	0.04±0.07 (0.08)	0.13±0.08 (0.15)	0.10±0.09 (0.14)
p = 5	n = 20	0.28±0.04	0.09±0.08 (0.12)	0.01±0.09 (0.10)	0.09±0.07 (0.12)	0.02±0.06 (0.06)	0.08±0.07 (0.11)	0.05±0.08 (0.09)
p = 5	n = 50	0.28±0.04	0.06±0.06 (0.09)	0.03±0.07 (0.07)	0.06±0.06 (0.08)	0.02±0.05 (0.06)	0.06±0.06 (0.08)	0.03±0.06 (0.07)

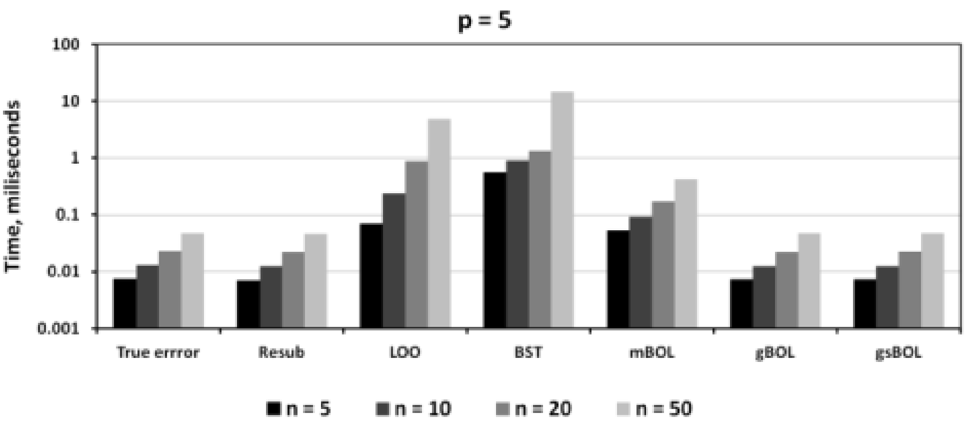


Figure 3. Representative timings in milliseconds for error estimators' performance depending on sample size for p = 5. Dataset – GSE13255

The calculation results also showed that the speed of the computation did not substantially depend on dimensionality (Figure 4).

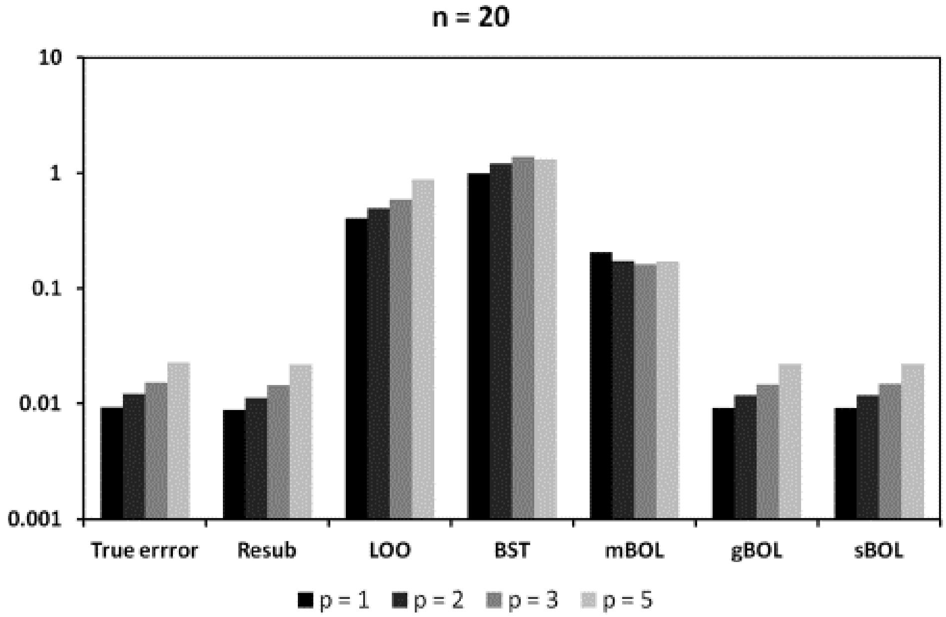


Figure 4. Representative timings in milliseconds for error estimators' performance depending on feature dimensionality for n = 20. Dataset – GSE13255

Similar results were obtained for the second dataset (GSE26712). The overall performance of the error estimators for this dataset is presented in Table 3 and Figure 5 and Figure 6.

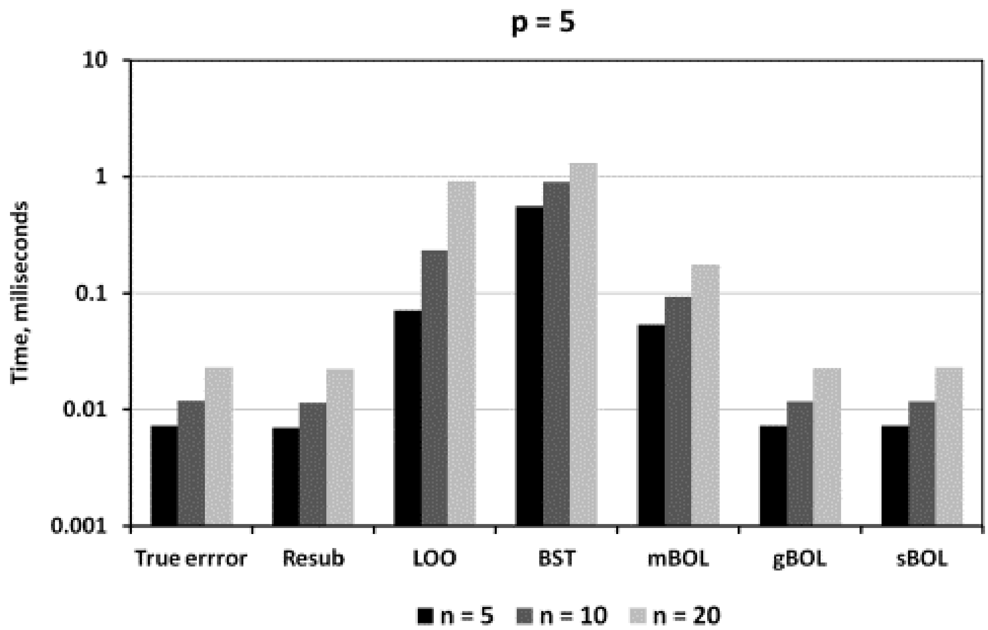


Figure 5. Representative timings in milliseconds for error estimators' performance depending on sample size for p = 5. Dataset – GSE26712

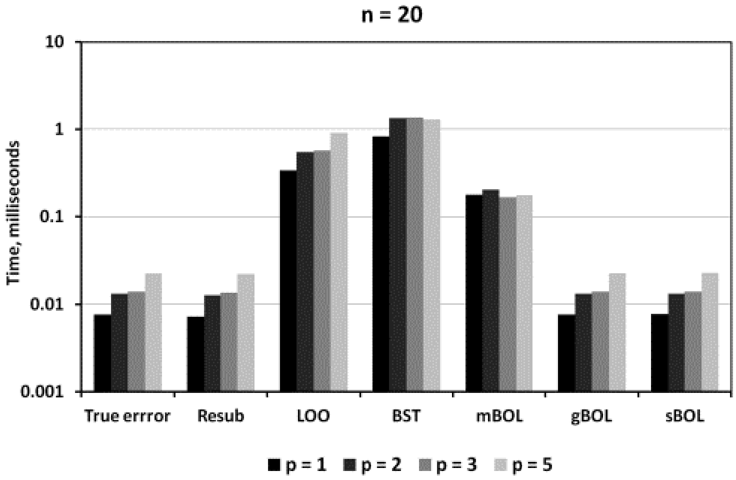


Figure 6. Representative timing in milliseconds for error estimators’ performance depending on feature dimensionality for n = 20. Dataset – GSE26712

Table 3. Performance of the error estimation algorithms in experiment 2 (GSE26712)

		True errror	Resub	LOO	BST	mBOL	gBOL	gsBOL
Fea-tures	Sample size	M ± SD	M_{diff} ± SD (RMS)					
p = 1	n = 5	0.44±0.09	0.07±0.13 (0.15)	-0.03±0.19 (0.19)	0.12±0.11 (0.16)	0.08±0.11 (0.13)	0.08±0.11 (0.14)	0.03±0.12 (0.12)
p = 1	n = 10	0.42±0.07	0.03±0.11 (0.11)	-0.02±0.14 (0.14)	0.06±0.09 (0.11)	0.04±0.10 (0.10)	0.03±0.10 (0.11)	0.01±0.10 (0.10)
p = 1	n = 20	0.40±0.05	-0.01±0.09 (0.09)	-0.03±0.10 (0.10)	0.01±0.08 (0.08)	-0.00±0.09 (0.09)	-0.01±0.09 (0.09)	-0.02±0.09 (0.09)
p = 2	n = 5	0.44±0.07	0.14±0.13 (0.19)	-0.02±0.19 (0.19)	0.22±0.09 (0.24)	0.10±0.10 (0.14)	0.13±0.11 (0.17)	0.09±0.12 (0.15)
p = 2	n = 10	0.42±0.06	0.08±0.11 (0.13)	-0.02±0.15 (0.15)	0.12±0.08 (0.15)	0.06±0.09 (0.11)	0.07±0.10 (0.12)	0.03±0.10 (0.11)
p = 2	n = 20	0.40±0.04	0.03±0.10 (0.10)	-0.01±0.11 (0.11)	0.06±0.08 (0.10)	0.03±0.08 (0.09)	0.03±0.09 (0.09)	0.00±0.09 (0.09)
p = 3	n = 5	0.44±0.07	0.19±0.14 (0.24)	-0.02±0.20 (0.21)	0.29±0.09 (0.30)	0.13±0.10 (0.16)	0.18±0.11 (0.22)	0.14±0.13 (0.19)
p = 3	n = 10	0.41±0.06	0.11±0.11 (0.15)	-0.02±0.15 (0.15)	0.17±0.08 (0.19)	0.08±0.08 (0.12)	0.10±0.10 (0.14)	0.07±0.10 (0.12)
p = 3	n = 20	0.40±0.05	0.05±0.10 (0.11)	-0.02±0.11 (0.12)	0.09±0.08 (0.12)	0.04±0.08 (0.09)	0.05±0.09 (0.10)	0.02±0.09 (0.09)
p = 5	n = 5	0.45±0.07	0.29±0.13 (0.32)	-0.02±0.21 (0.21)	0.37±0.08 (0.38)	0.16±0.09 (0.18)	0.27±0.11 (0.29)	0.24±0.12 (0.27)
p = 5	n = 10	0.42±0.06	0.18±0.11 (0.21)	-0.02±0.16 (0.16)	0.26±0.08 (0.27)	0.11±0.08 (0.13)	0.17±0.09 (0.19)	0.13±0.10 (0.17)
p = 5	n = 20	0.39±0.06	0.09±0.10 (0.13)	-0.01±0.12 (0.12)	0.10±0.08 (0.13)	0.06±0.08 (0.10)	0.09±0.09 (0.13)	0.06±0.09 (0.11)

Complete raw data for both experiments are available at http://www.molbiol.sci.am/big/jbc/Arakelyan_JBC_src.zip (Performance_GSE13255.xls and Performance_GSE26712.xls in archive file).

Discussion

In this paper we have compared the performance of our geometrical bolstered and semi-bolstered error estimation algorithms with their original counterparts [Braga-Neto, 2004], as well as several popular error estimation techniques [Devroye, 1996; Lachenbrucha, 1968; Efron, 1983]. Bolstered error estimation is used for estimation of classification true error and has several advantages over other contemporary algorithms in terms of accuracy and speed. For general cases of classification, the bolstered error is estimated by Monte-Carlo integration, but in specific cases it can be computed exactly by solving the integral-containing equations described in [Braga-Neto, 2004]. Here, we propose a geometric solution to the bolstered error estimation, namely gBOL and gsBOL algorithms, specifically designed for the case of linear classification and spherical bolstering.

There are two key features that are characteristic to a good error estimator: accuracy and speed. In terms of accuracy, both gBOL and gsBOL perform comparable to other commonly used error estimation techniques, such as LOO and BST, as well as the mBOL algorithm, proposed in [Braga-Neto, 2004]. Moreover, gBOL and gsBOL are at least 10-350 times faster than LOO and BST and 5-10 times faster than mBOL depending on the sample size. The speed issue is very important when performing feature selection from extremely high numbers of features. For example, error estimation with LOO for 5-gene feature set takes almost 1 second, while gBOL or gsBOL do the same with comparable accuracy in 0.05 second. This means that for 10000 feature sets (which is a quite common situation, e.g. in microarray gene expression data analysis), a user will spend about 3 hours with LOO, while with gBOL or gsBOL – only about 10 minutes. This advantage will allow for saving time and computing resources, and concentrating more on data analysis results interpretation.

We acknowledge the fact that for more complex classifiers and other kernel types, our approach might not be fully applicable; however, we also recognize that in biomedical research, linear classification is among the most frequently used classification techniques [Schölkopf, 2004; Tarca, 2007; Ben-Hur, 2008]. Thus, gBOL and gsBOL implementation of bolstered error estimation is an optimal choice in the case of linear classifiers. We believe that gBOL and gsBOL will find their users at least in the fields of genomics and structural bioinformatics, where linear classification is being actively used [Lee, 2011; Holloway, 2007; Yang, 2004].

Finally, gBol and gsBOL algorithms do not include any special MATLAB commands or functions and can be easily implemented in any other programming language, like C/C++, R, FORTRAN, and Java.

Conclusions

Taken together, gBOL and gsBOL are very fast error estimation techniques characterized by accuracy comparable with LOO and with lower variance. These algorithms are useful for analyzing extremely large numbers of features and may find their application in various fields of - *omics* data analysis.

Acknowledgments

This study was supported by Armenian National Science and Education foundation (NS molbio-3507 to A.A.).

Bibliography

[Allison, 2006] D. B. Allison, X. Cui, G.P. Page, and M. Sabripour, “Microarray Data Analysis: from Disarray to Consolidation and Consensus,” Nature Reviews Genetics, vol. 7, no. 1, pp. 55–65, Jan 2006.

[Ben-Hur, 2008] A. Ben-Hur, C.S. Ong, S. Sonnenburg, B. Schölkopf, and G. Rätsch, “Support Vector Machines and Kernels for Computational Biology,” PLoS Computational Biology, vol. 4, no. 10, e1000173, Oct 2008.

[Bonome, 2008] T. Bonome, D.A. Levine, J. Shih, M. Randonovich, C.A. Pise-Masison, F. Bogomolniy, L. Ozbun, J. Brady, J.C. Barrett, J. Boyd, and M.J. Birrer, “A Gene Signature Predicting for Survival in Suboptimally Debulked Patients with Ovarian Cancer,” Cancer Research, vol. 68, no. 13, pp. 5478–5486, Jul 2008.

[Braga-Neto, 2004] U. M. Braga-Neto, and E.R. Dougherty, “Bolstered Error Estimation,” Pattern Recognition, vol. 37, no. 6, pp. 1267-1281, June 2004.

[Cai, 2003] C. Z. Cai, W.L. Wang, L.Z. Sun, and Y.Z Chen, “Protein Func-tion Classification via Support Vector Machine Approach,” Mathematical Biosciences, vol. 185, no. 2, pp. 111-22, Oct 2003.

[Choi, 2011] S. Choi, and K. Han, “Prediction of RNA-binding Amino Acids from Protein and RNA Sequences,” BMC Bioinformatics, 12 Supple 13:S7, Nov 2011, doi: 10.1186/1471-2105-12-S13-S7.

[Cui, 2012] G. Cui, C. Fang, and K. Han, “Prediction of Protein-protein Interactions between Viruses and Human by an SVM Model,” BMC Bioinformatics, 13 Suppl 7:S5, May 2012, doi: 10.1186/1471-2105-13-S7-S5.

[Devroye, 1996] L. Devroye, L. Györfi, and G. Lugosi, “A Probabilistic Theory of Pattern Recognition”, Springer, New-York, 1996.

[Dougherty, 2010] E. R. Dougherty, C. Sima, J. Hua, B. Hanczar, and U.M. Braga-Neto, “Performance of Error Estimators for Classification,” Current Bioinformatics, vol. 5, no. 1, 53-67, March, 2010.

[Efron, 1994] B. Efron, and R.Tibshirani, “An Introduction to the Bootstrap”, Chapman & Hall, New York, 1994.

[Efron, 1983] B. Efron, “Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation,” Journal of the American Statistical Association, vol. 78, no. 382, pp. 316-333, 1983.

[Golub, 1999] T. R. Golub, D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, J.R. Downing, M.A. Caligiuri, C.D. Bloomfield, and E.S. Lander, “Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring,”Science, vol. 286, no. 5439, pp. 531-7, Oct 1999.

[Hastie, 2009] T. Hastie, R.Tibshirani, and J. Friedman, “The Elements of Statistical Learning”, Second Edition: Data Mining, Inference, and Prediction, Springer, New York, 2009.

[Holloway, 2007] D.T. Holloway, M. Kon, and C. Delisi, “Machine Learning for Regulatory Analysis and Transcription Factor Target Prediction in Yeast,” Systems and Synthetic Biology, vol. 1, no. 1, pp. 25-46, Mar 2007.

[Lachenbrucha, 1968] P.A. Lachenbrucha, and M.R. Mickey, “Estimation of Error Rates in Discriminant Analysis”, Technometrics, vol. 10, no. 1, pp. 1-11, 1968

[Lee, 2011] D. Lee, R. Karchin, and M.A. Beer, “Discriminative Prediction of Mammalian Enhancers from DNA Sequence,” Genome Research, vol. 21, no. 12, pp. 2167-80, Dec 2011.

[Li, 2011] S. Li, “Concise Formulas for the Area and Volume of a Hyperspherical Cap,” Journal of Mathematics & Statistics, vol. 4, no. 1, pp. 66-70, 2011.

[Maulik, 2013] U. Maulik, A. Mukhopadhyay, and D. Chakraborty, “Gene-Expression-Based Cancer Subtypes Prediction Through Feature Selection and Transductive SVM,” IEEE Transactions on Bio-Medical Engineering, vol. 60, no. 4, pp. 1111-7, Apr 2013.

[Nguyen, 2011] M.N. Nguyen, J.M. Zurada, and J.C. Rajapakse, “Toward Better Understanding of Protein Secondary Structure: Extracting Prediction Rules,” IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 8, no. 3, pp. 858-64, May 2011.

[Nolan, 1997] J.R. Nolan, “Estimating the True Performance Of Classifica-tion-Based NLP Technology,” Proc. From Research to Commercial Applications: Making NLP Work in Practice, pp.23-28, 1997.

[Nolan, 1997] T.R. Nolan, “DISXPERT: A Rule-Based Vocational Rehabilitation Risk Assessment System,” Expert Systems with Applications, vol. 12, no. 4, pp. 465-472, May 1997.

[Okun, 2011] O. Okun, Feature Selection and Ensemble Methods for Bioinformatics: Algorithmic Classification and Implementations. Medical Information Science Reference, Hershey, 2011.

[Sayes, 2007] Y.Saeys, I. Inza, and P. Larrañaga, "A Review of Feature Selection Techniques in Bioinformatics," Bioinformatics, vol. 23, no. 19, pp. 2507-17, Oct 2007.

[Schölkopf, 2004] B. Schölkopf, K.Tsuda, and J.P. Vert, Kernel methods in computational biology. MIT Press, Cambridge, 2004.

[Showe, 2009] M.K. Showe, A. Vachani, A.V. Kossenkov, M. Yousef, C. Nichols, E.V. Nikonova, C. Chang, J. Kucharczuk, B. Tran, E. Wakeam, T.A. Yie, D. Speicher, W.N. Rom, S. Albelda, and L.C. Showe, "Gene Expression Profiles in Peripheral Blood Mononuclear Cells Can Distinguish Patients with Non-small Cell Lung Cancer from Patients with Nonmalignant Lung Disease," Cancer Research, vol. 69, no. 24, pp. 9202-10, Dec 2009.

[Tarca, 2007] A. L. Tarca, V.J. Carey, X.W. Chen, R. Romero, and S. Drăghici, "Machine Learning and its Applications to Biology," PLoS Computational Biology, vol. 3, no. 6, pp. e116, Jun 2007.

[Vapnik, 1995] V. Vapnik, "The Nature of Statistical Learning Theory", Springer, New York, 1995.

[Yang, 2004] Z. R. Yang, "Biological Applications of Support Vector Ma-chines," Briefings in bioinformatics, vol. 5, no. 4, pp. 328-38, Dec 2004.

Authors' Information

Arsen Arakelyan *– Bioinformatics group of the Institute of Molecular Biology NAS RA, Laboratory of Information Biology of the the Institute of Molecular Biology and Institute for Informatics and Automation Problems NAS RA, 7 Hasratyan St, 0014, Yerevan, Armenia, email: aarakelyan@sci.am*

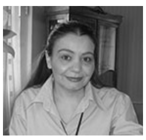
Major Fields of Scientific Research: Algorithm development foe gene expression analysis, gene network analysis and modeling

Lilit Nersisyan *- Bioinformatics group of the Institute of Molecular Biology NAS RA, 7 Hasratyan St, 0014, Yerevan, Armenia, email: l_nersisyan@mb.sci.am*

Major Fields of Scientific Research: Algorithm development foe gene expression analysis, gene network analysis and modeling

Aram Gevorgyan *- Bioinformatics group of the Institute of Molecular Biology NAS RA, 7 Hasratyan St, 0014, Yerevan, Armenia, email: imb@sci.am*

Major Fields of Scientific Research: Algorithm development foe gene expression analysis, gene network analysis and modeling



Anna Boyajyan *– Department of Applied Molecular Biology of the Institute of Molecular Biology NAS RA, Laboratory of Information Biology of the Institute of Molecular Biology and Institute for Informatics and Automation Problems NAS RA 7 Hasratyan St, 0014, Yerevan, Armenia, email: aboyajyan@sci.am*

Major Fields of Scientific Research: Genomics & Immunomics

THE MANAGEMENT OF PATIENT INFORMATION IN POLISH HEALTH CARE SYSTEM

Anna Sołtysik-Piorunkiewicz

Abstract: *The purpose of the paper is to comprehensively characterize the implementation and the current use of electronic health records (EHRs) in Poland. The paper presents the conception of the management of patient information in Polish health care system. In this paper author presents the assumptions of the strategies of implementing IT systems in the health protection sector in Poland. It contains a review of IT systems dedicated for health care centers. The author described a system of electronic medical records, taking into account governmental plans for implementing the Health Insurance Card and Polish.ID systems, as well as the Electronic Verification of Beneficiary Entitlements (EVBE) system being implemented by Polish government. The project of the EVBE system is a newest part of Polish health care system. The paper presents also an example of electronic European Health Insurance Card system to management of patient information in Polish health care system supported by the EVBE. European Health Insurance Card was created as a response to the necessity of verification of a patient in the European Union and in the countries of the Schengen area. The aim of the article is to describe the difference between the public electronic patient record system in Poland and the private IT systems dedicated for health care centres. The author described a system of electronic medical report, taking into account governmental plans for implementing the Health Insurance Card and pl.ID systems, as well as e-Prescription system being implemented by the Centre for Health Information Systems.*

Keywords*: IT systems in health care, electronic health insurance card, health informatics, electronic medical records.*

ACM Classification Keywords*: K6: management of computing and information systems*

Introduction

The aim of management of patient information in health care system in such facilities as health care centers (hospitals, clinics), pharmacy's shops and others, is to improve the effectiveness of spending public funds on health care. Such systems should be characterized by integration and capability of extension [Sołtysik-Piorunkiewicz, 2012].

IT systems, which are used to streamline, improve the functioning of health care, are the area of interest as well in Poland [Kawiorska, 2004; Pankowska 2004], as in many countries across the world and Europe, e.g. Germany [Kalkulation von Fallkosten, 2002], Switzerland [Thatcher, 2013], Netherlands [Shekelle et al, 2006]. The transformation, due to implementation of IT systems, into a uniform, integrated and flexible structure, is undoubtedly beneficial for the safety and comfort of patients and streamlining activities in a health care system [Kaplan & Harris-Salamone, 2005]. Benefits from implementing IT systems fulfil the expectations of patients and health services providers, as well as improving the whole process of decision making by participants of the health care system. The research of United States academic medical centers reported generally positive attitudes towards using the electronic health record (EHR) as a structured, distributed documentation systems that differ from paper charts [Han & Lopp, 2013] in the ambulatory setting and significant concerns about the potential impact of the EHR on their ability to conduct the doctor-patient encounter [Rouf et al, 2008]. However the clinical documentation, an essential process within electronic health records (EHRs), takes a significant amount of clinician time. There are luck of methods to optimize documentation ways to deliver effective health care [Pollard et al, 2013].

The results of a large-scale statewide analysis in 2006 in United States showed the adoption of health information technologies in physician's offices [Menachemi & Brooks 2006; Menachemi et al, 2006]. They discovered that the use of quality enhancing technologies such as PDAs, use of e-mail with patients and EHR was less common. But the state of adoption of health information technologies in physician's offices is still changing. The American Medical Information Association (AMIA) recommendations is based on research and publication, best practices, advocacy, education, certification, databases and knowledge integration. With the United States joining other countries from Europe, e.g. The Great Britain, in national efforts to reap the many benefits that use of health information technology can bring for health care quality and savings, to minimize complexity and difficulties of implementing even smaller-scale systems [Kaplan & Harris-Salamone, 2005; Kaplan & Harris-Salamone, 2009]. The last research about factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy, based on cooperation between RAND Health, a division of the RAND Corporation and AMA (American Medical Association), was published in 2013 [Friedberg et al, 2013]. A major factor limiting efficiency and quality gains from clinical information technologies is the lack of full use by the clinicians [Kralewski et al, 2008].

Projects for implementing IT systems in health care for management of patient information are complex and complicated undertakings. These processes should be carried out in stages. To achieve best effects, individual stages should be implemented efficiently and consistently. Moreover, appropriate legislative changes are required and – importantly – persuading decision makers about necessity of such solutions.

Any IT systems implemented in health care must comply with legal standards in force and be adaptable to changes made to the national law. Due to the mass scale of health care, they must also be efficient both for patient and health protection system [Sołtysik-Piorunkiewicz, 2012].

There are many benefits of health information technology that can be bring more health care quality and savings, sobering reports recall the complexity and difficulties of implementing even smaller-scale systems [Kaplan & Harris-Salamone, 2009]. The ability of EHRs to improve the quality of care in ambulatory care settings was demonstrated improvements in provider performance when clinical information management and decision support tools were made available within an EHR system [Shekelle et al, 2006]. The developers, implementers and certifiers of EHRs should focus on increasing the adoption of robust EHR systems and increasing the use of specific features rather than simply aiming to deploy an EHR regardless of functionality to maximize health care quality [Poon et al, 2010].

Over the past several years, Sittig and others have carried out an extensive qualitative research program focusing on the barriers and facilitators to successful adoption and use of advanced, state-of-the-art clinical information systems based on commercially available EHR vendors and the internally developed EHRs. They concluded that if the well-designed commercially-available systems are coupled with the other key socio-technical concepts required for safe and effective EHR implementation and use, and organizations have access to implementable clinical knowledge, the transformation of the healthcare enterprise that so many have predicted, is achievable using commercially-available, state-of-the-art EHRs [Sittig et al, 2011].

Accelerating the adoption of IT health care systems will require greater public-private partnerships, new policies to address the misalignment of financial incentives, and a more robust evidence base regarding IT implementation [Goldzweig et al, 2009].

The directions of implementing IT systems in Polish health care

EHR implementation is essential to improving patient safety but is still highly heterogeneous across health care systems and providers, and this heterogeneity leads to equally variable implications for patient safety [Sittig & Singh, 2012].

The directions of implementing IT systems in Polish health care are now concentrated mainly on putting into effect the assumptions of the European Commission concerning e-Health. The main issues presented in the „e-Health Poland" plan which require implementation by 2015 include the following [Directions of inf., 2009]:

1. Ensuring citizens easier access to health care information.
2. Improving effectiveness of the health care system with regard to electronic flow of information.
3. Creating procedures and guidelines, gathering and giving access to good practices to improve management of a health care centre thanks to implementing information and communication systems.
4. Modernizing the system of medical information to analyze the demand for provided health services.
5. Practical realization of the development of IT solutions in protection of health in line with the guidelines the European Commission which will allow the Republic of Poland to be included into the area of interpretational Electronic Health Record.

Currently, due to accepting for implementation Programmes financed from structural measures, the following activities are of key importance:

- Implementation of the Programme of Health Protection Informatization;
- Creating the conditions for the development of health protection e-services – especially telemedical systems (teleconsultations, telemonitoring, online patients' registration) e-prescriptions and electronic health records, which will be linked with a new identity card [Sołtysik-Piorunkiewicz, 2012].

Review of Polish health protection IT systems

Health insurance card

Electronic health insurance card was implemented in the Silesia voivodeship about 10 years ago as part of works connected with informatization of Silesia voivodeship department of National Health Fund. Electronic health insurance card is used to verify insurance status of an authorized card holder in the system of Silesia voivodeship department of National Health Fund. It also provides personal data and is used for authorization of the services provided as part of performing a contract with Silesia voivodeship department of National Health Fund [Karta Ubezpieczenia Zdrowotnego, 2012]. This card may be issued to an insured person with a Universal Electronic System for Registration of the Population national identification number who can prove his/her residence on the territory of the Silesia voivodeship.

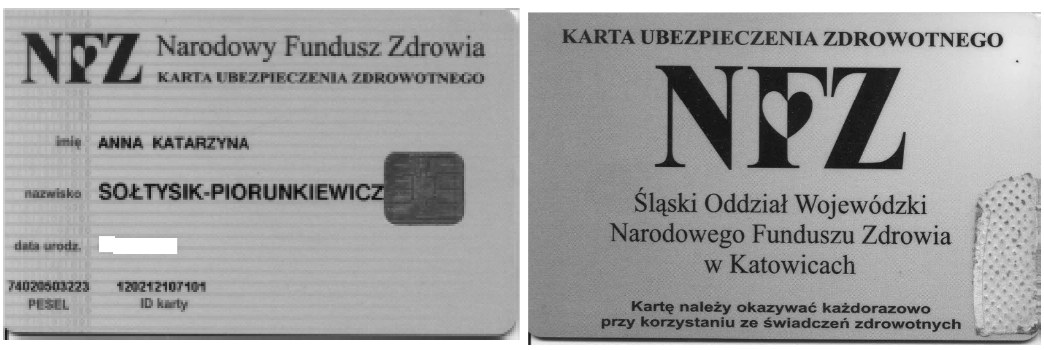


Figure 1. Electronic health insurance card of Silesia voivodeship department of National Health Fund

Since 2004, it has been planned in Poland to implement country-wide Health Insurance Card modelled on the Silesian card. To continue the works ordered by Health Minister on February 8, 2005, a Group was set up to draw up a strategy for the development of a medical information system in health protection and prepare a conception of implementing European Health Insurance Card and Health Insurance Card. The tasks of this group included drawing up a strategy for the development of a medical information system in health protection and preparing a conception of implementing European Health Insurance Card and Health Insurance Card. It was decided that representatives of Health Ministry would be involved in the works of a Group for developing a conception and design of electronic Health Insurance Card (e-HIC), electronic Medical Services Register (e-MSR) and programmes of their implementation, set up by order of the President of the National Health Fund No 40/2004 of 25 November 2004 for the purpose of preparing a functional conception, scope of application and design of e-HIC, linked with European Health Insurance Card, and a conception and design of a system for electronic registration and medical services monitoring (e-MSR) and programmes (strategy, plan, schedule) for implementing e-HIC and e-MSR [Reply of the Secretary, 2013]. Up to now, the implementation has not been completed.

Another proposal of implementing electronic medical report was using the function of a health insurance card in an electronic identity card as part of the implementation of the programme MSR II. This project was described in a publication on the connection between Health Insurance Card and pl.ID. Its basic aim is to ensure the verification of the parties, place and sequence of a medical transaction (patient and professionals) by means of a cryptographic secret carrier. This project was based on the use of a crypto processor card (Health Insurance Card, Professional's Card) in an environment that was safe for creation of electronic signature. It was also assumed that pl.ID would constitute an electronic document which might be used for verifying a person (including limited identification), creating personal and qualified signatures, and entitles to cross the borders of the countries united by the Schengen Agreement, as well as serves the function of HIC. As a result of a decision by National Council of Ministers of December 2009, the project was linked with pl.ID, in which there was a separate space for HIC. The National Health Fund was developing and handing to the Ministry of the Interior and Administration the document „Technical and functional requirements for Health Insurance Card, HIC application and their environment".

Electronic European Health Insurance Card as a part of a Polish health care system for management of patient information

One of the elements of a Polish health care system in managing patient information is electronic European Health Insurance Card. It was created as a response to the necessity of verification of a patient in the European Union and in the countries of the Schengen area.

Figure 2. European Health Insurance Card

Currently the project of Health Insurance Card for the patient developed in connection with the pl.ID system is suspended in Poland. However, the work on the system of European Health Insurance Card is still underway due to the development of a new conception of the system of patient identification in the countries of European Free Trade Association [Departament Współpracy Międzynarodowej Centrali NFZ, 2012]. The most important part of implementation of pl.ID in polish National Registers System is now in the stage of consultation and analysis of future functionalities [National Registers System, 2013].

This card may be issued to an insured person with a national identification number Universal Electronic System for Registration of the Population who can prove his/her residence on the territory of the Silesia province. To continue the works ordered by the Health Minister on February 8, 2005, a Group was set up to draw up a strategy for the development of a medical information system in health care, and prepare a conception of implementing European Health Insurance Card and Health Insurance Card. The tasks of this group included drawing up a strategy for the development of a medical information system in health protection and preparing a conception of implementing European Health Insurance Card and Health Insurance Card. The main aim of the implementation of this solution was to [Ujejski, 2011]:

- Facilitate the process of registration and confirming the right to health services (effectiveness);
- Improve the reliability of accounting data sent to the National Health Fund (integrity and indisputability);
- Secure access to data (confidentiality);
- Increase the chance of a rescue in emergencies;
- Satisfaction of the entitled, possibility of getting remote access to own data;
- Increase the effectiveness of medical centres (automation of activities);
- Decrease the number of frauds, mainly in the area of accounting for services that have not been provided;
- Improve the quality of data.

Ultimately, health insurance card should be an element of the information system in the health protection system in Poland. However, using such a card entails access to sensitive data, thus one of basic requirements is ensuring security of the information system used in medical services. It should also be able to integrate with the medical services register and medication register. It may also be an instrument for checking insurance as part of processing European Health Insurance Card instead of the National Health Fund slip of the monthly report for the insured person. Electronic medical report should be integrated with the information system for pharmacies and health care centres as well as the electronic system for prescriptions and electronic system for medical appointment referrals.

The Electronic Verification of Beneficiary Entitlements for management of patient information

At the same time, the Electronic Verification of Beneficiary Entitlements project is being developed in Poland, i.e. Electronic Verification of Beneficiary Entitlements, in compliance with the law on the healthcare services financed from the public funds. From January 1, 2013 patients visiting doctors should present Universal Electronic System for Registration of the Population number and identity card, driving licence or passport, in order to confirm their entitlement to receiving healthcare from public funds [Do lekarza, 2012]. The tests of the system were conducted from October 15 to the end of 2012, and its aim is to introduce orderliness into the system of confirming the entitlement to receive healthcare services. It also takes from doctors the responsibility for checking whether the patient is entitled. The project was created from the cooperation of the Ministry of Administration and Digitalization with the Ministry of Health, National Health Fund (Polish ZUS), Social Insurance Institution and

Agricultural Social Insurance Fund, and was consulted with, among other things, the Chief Inspectorate of Personal Data Protection.

Thanks to information exchange among the Central List of the Insured of the National Health Fund and Agricultural Social Insurance Fund registers, it will be possible to check online in a hospital or clinic reception whether the patient is insured. There will be no need for the National Health Fund slip of the monthly report for the insured person (although it still will be valid, just like the pensioner identity card and other documents entitling for receiving services).

In the case when the system did not confirm the entitlement to receiving the services (e.g. when the employer did not register the employee for insurance), the statement of the insured person will suffice (art. 7 and 8 of the relevant law).

The introduction of the Electronic Verification of Beneficiary Entitlements system made the procedure of issuing the National Health Fund slip of the monthly report for the insured person more efficient. The citizen doesn't have to produce the monthly report for the insured person slip in the National Health Fund institution, instead he/she can show identity document.

Integrated IT systems for comprehensive service of health care centres and settlements with NHF provided by Asseco Poland SA

Depending on the needs of a healthcare provider with regard to medical services, Asseco Poland SA offers various IT systems for implementation. These are:

- InfoMedica [Infoklient bi, 2014];
- mMedica in versions PS, PS +, Standard and Standard + [Mmedica, 2014];
- Hipokrates;
- SolMed.

The InfoMedica system consists of several modules. These are, among others: medical systems, and administration and management systems. The medical systems include the following packages:

- Package: Hospital;
- Package: Clinic Pro;
- Package: Diagnostics;
- Dialysis unit;
- Workplace infections;
- Package: Laboratory with microbiology;
- Ambulance transport;
- Package: eMedica portal which includes eKontrahent and ePacjent packages.

InfoMedica medical systems are installed in admission rooms, in hospital wards, in diagnostic laboratories, laboratories, in operating blocks, in doctor's surgeries and treatment rooms, in clinics and outpatient clinics, in hospital chemists and medical statistic departments. They enable an efficient gathering and distribution of all medical information connected with the history of treatment of each patient, from his/her admission to the hospital until the treatment is finished. They help doctors to assess a patient's health condition, make it easier to access the data on the health condition of each patient being treated, print all forms used in the treatment process, ensure obligatory statistical reporting for the National Health Federation as well as medical statistics institutions and centres. They streamline the organization of treatment process. The dedicated Clinic Pro package supports the work of medium-sized and large out-patient health care centres. It has got modules to be installed at reception desks and in doctor's surgeries. It offers specialized functionalities dedicated for supporting occupational medicine surgeries, dentist's surgeries, and rehabilitation centres. Due to full integration, doctors have access in their surgeries to a patient's whole medical documentation, including the history of hospitalization as well as examinations and medical procedures carried out [Informklient systemy medyczne, 2012].

The administration and management systems include the following packages:

- Package: finances and accounting;
- Package: Treatment Costs Bill;
- Package: Sale Service;
- Package: Budgeting – Controlling;
- Package: Managing the trade of medicine and materials;
- Package: Managing Fixed Assets;
- Package: Human Resources and Payroll Service;

The administration and management systems of the InfoMedica package are responsible for gathering and processing all information connected with economic events in a hospital. They are installed in accounting department, Human Resources department, salary calculation, and other departments of a hospital's administration. They are used to maintain comprehensive bookkeeping, manage finances, carry out management accounting, prepare price lists of medical services and offers for, among others, the National Health Federation, and other health care payers, develop plans for the sale of medical services, monitor contracts and agreements, and receipts from their fulfilment, maintain a detailed costing of current activity, calculate costs of patients' treatment, perform economic predictions. All the applications of the InfoMedica package are integrated with each other so as to ensure a flow of information between relevant organizational units of a hospital in which IT system was implemented [Infoklient systemy administracyjno zarządcze, 2012].

The system of electronic prescription e-Prescription

The e-Prescription system is one of the first projects of building a modern IT system in health protection. It is being implemented under the project „Electronic Platform for Gathering, Analysing and Sharing digital resources about medical events", which is part of the country-wide Programme for Health Protection Informatization. The entity responsible for implementing the electronic prescription system is the Centre for Health Information Systems, set up by the Health Minister [Poznaj, 2012]. All the information gathered in the e-Prescription system is protected in accordance with security standards in force. The information in the e-Prescription system is exchanged using encrypted SSL protocol (Secure Sockets Layer), which is a widely used data transmission standard. This allows for the communication in the e-Prescription system to take place only between entitled participants of the prototype and an unauthorized access to data is prevented. Moreover, cryptographic techniques are used to verify and ensure information consistency. The technique of encryption is based on process of transforming plain text or data into cipher data that cannot be read by anyone other then the sender and the intended receiver [Laudon & Laudon, 2011]. Usage of encryption is necessary to protect digital information that are stored, transferred, or sent over the Internet. The capability to generate secure session is built into Internet client browser software and servers. SSL is designed to establish a secure connection between both two computers of the client and the server.

The e-Prescription system allows electronic prescription to function parallel with its formal counterpart in the form of a paper prescription. The use of electronic prescription along with a paper prescription does not disturb the existing model of prescriptions circulation which is based only on paper prescriptions. Currently, due to the lack of appropriate regulations regarding the functioning of prescriptions only in an electronic form, an electronic document does not constitute a prescription in the light of law.

Writing a prescription in an electronic form takes place by means of a doctor's medical computer program, and in the case when a doctor does not have internal software in his/her surgery, communication with the e-Prescription system occurs through an on-line application.

Pharmacists work with the e-Prescription system directly through their chemist's program, and can also dispense electronic prescriptions through a dedicated on-line application.

Patients get access to information about the history of their pharmacotherapy through the e-Prescription Internet Account.

The prototype of the e-Prescription system was launched in mid-March 2011, but the actual processing of prescriptions using this system started on 18 April 2011. The delays were connected with the need to install and launch the system in the entities that signed a cooperation agreement with the Centre for Health Information Systems, and necessity of providing training to medical personnel and pharmacists on how to operate the system.

The e-Prescription system supported 20 facilities, i.e. 2 clinics, 2 doctor's surgeries and 16 chemist's shops. This implementation can be regarded as a pilot implementation across the country in 2012 [Sołtysik-Piorunkiewicz, 2012]. However, for the system to be successfully implemented across the whole country, electronic prescription should be regulated by new legal regulations, and in the whole country there should be access to the public database of medicines, patient's insurance card and a tool for doctors authentication.

Systems for pharmacy's shops and clinics provided by Kamsoft SA

One of the companies which provide IT systems for health protection sector is Kamsoft SA based in Katowice. The company is developing a lot of projects in medicine and pharmacy, from systems for supporting hospitals (KS-MEDIS), clinics (KS-SOMED) or dentist's clinics (KS-KST), through data security systems in medicine and pharmacy (KS-BDO) and a database of medicine and health protection means (KS-BLOZ), to national system of health protection. The company also implements systems supporting chemist's shops: Integrated System for Managing a Network of Chemist's Shops (KS-ZSA) and IT System for Supporting the Handling of a Chemist's Shop (KS-AOW) and systems for pharmaceutical wholesale outlets (KS-EWD).

The Integrated IT System for Servicing a Clinic (KS-SOMED) is a multi-module tool for supporting the work of medium-sized and large specialist clinics. The system is characterized by an extended functionality which enables the processing of the most important organizational issues:

- Managing doctor's appointments schedule;
- Gathering of medical data;
- Handling of financial settlements with payers.

The NHPS National Health Protection System portal is a free of charge educational and informational platform for all representatives of the health protection market and patients. It enables communication between a patient, doctor and pharmacist, the use of various health services and access to preventive health programs. The central element of the NHPS system is patient and his/her family. Each patient is represented in the system by an Individual Health Account. The Patient Service enables performing various operations on the health account. They include viewing data gathered in the system and entering certain information by a patient himself. This information may be subsequently used by doctors and medical personnel to have a better knowledge about patient's health, and thus it may lead to providing the patient with a better health protection [Strug, 2009].

The Electronic Data Exchange system (KS-EWD) enables direct communication between a chemist's shop and a wholesale outlet by means of the Internet. The unique features include interactive checking availability of goods in a wholesale outlet, together with the option of immediate ordering of the goods on special offer. Due to integration of the KS-AOW and KS-EWD systems, the system designed for chemist's shops includes features ensuring data exchange between a chemist's shop and a wholesale outlet.

The IT System for Supporting the Handling of a Pharmacy's Shop (KS-AOW) is a comprehensive system for supporting the work of chemist's shops, working under the Windows environment. In its work, the system uses SQL databases and Internet technologies. Particular emphasis was placed in the system on an efficient and fast operation. Also, the solutions applied in the system allow using needle printers and their work with the system in a character mode. The KS-AOW system consists of a dozen interrelated modules, which ensure a wide range of possibilities. The features included in the modules comprising the System allow checking sale, orders and purchases, manage the warehouse, create lists, analyses and perform accounting. The KS-AOW system complies with the regulations of the Health Ministry, Finance Ministry, National Health Fund and other entitled institutions and is constantly modernized and adapted to changing regulations.

Integrated System for Managing a Network of Pharmacy's Shops (KS-ZSA) is a comprehensive system supporting monitoring and managing the work of a network of chemist's shops, working under the Windows environment. The central management of chemist's shops enables an optimal warehouse management, rationalization of purchases and a better use of the economies of scale during purchases. The system consists of a dozen interrelated modules which ensure a wide range of possibilities. The features included in the modules allow to place orders and make purchases, manage the warehouse, create lists, analyses as well as accounting for individual chemist's shops. Management of chemist's shops, by analysis and purchase specialist, supported by the KS ZSA system, allows relieving the staff of chemist's shops of many tasks and using the time saved for increasing professionalism of patient care and assistance – which is a more and more important element of competing on the pharmaceutical market [Direction, 2011].

Discussion

There are some lacks of information health care systems in Poland:

- Certain data redundancy – similar data is stored in different registers, kept by two different entities, e.g. Central List of Insured Persons and Central Register of Insured Persons (The National Health Federation and Health Insurance Company);
- Incoherence of data between different registers – the same data in different registers may vary (e.g. change of address is not automatically updated in all registers in which the address appeared);
- Lack of cooperation with reference registers – generally, registers do not refer to base reference registers;
- Lack of cooperation between registers in health protection – registers in health protection do not use source information which already exists in other health protection registers;
- Lack of a uniform data model – registers and databases in health protection do not use a uniform data model;
- Lack of structure and relationships between registers in health protection.

Currently, newer and newer systems for pharmacists' shops are being implemented, i.e. EuroMedica [Apteka Euromedica, 2012], based on modern Internet technologies. Clinics implement InfoMedica or OSOZ systems (Table 1). However, it seems necessary to implement such IT systems in health protection for management of patient information in Polish health care system e.g. a Health Patient Information System [System, 2013] which will ensure [Directions, 2011]:

1. Creation of information conditions that will allow taking long-term optimal decisions in health policy, irrespective of the adopted organizational model of health protection and principles of financing it.
2. Creation of a stable information system in health protection, characterized by a flexible approach to the organization of the system of health protection resources, including a model of financing services from public funds, and resilience to disturbance in data gathering and archiving, caused by system changes in health protection.
3. Decreasing the information gap in the health protection sector that makes it impossible to build an optimal model of health protection.
4. Organizing an existing system for information gathering, processing and analysing (e.g. in Business Intelligence systems).
5. Building a system ensuring electronic communication and possibility of exchanging documents and reports between health protection entities and the proprietary body.

Table 1. Manufacturers, name of health care systems and features of them

Producers	Name System/Module	Charcteristics of Features
Asseco Poland SA	InfoMedica	Package programs modular registrants health benefits (part white) and administrative and economic events (gray part) and the settlement of the contributors. It has elements of data analysis and business decision support.
	Hipokrates	It covers an area medical (white), administrative (gray) and management support. Implemented electronic medical history. One of the first system in Poland HIS.
	Solmed	The system is dedicated for smaller entities supporting the movement of patients, the primary drug economy and the most important elements of the branch operation and administration.
Gabos Software Sp. z o.o.	Mediqus 3	A comprehensive service system of health care facilities . Supports medical part, including issues related to the management and operation of both hospitals and outpatient units.
Kamsoft SA	KS-MEDIS	HIS system that supports part of the white and gray hospital.
	KS-SOLAB	Laboratory system, serving both small laboratories and hospital units.
	KS-SOMED	The outpatient service.
	KS-ZSA	Pharmacy network management system.
	KS-AOW	Support system for pharmacy.
	KS-KST	A dedicated system for dental treatment.

Conclusion

The paper discusses the difference between the public electronic patient record system in Poland and the private IT systems dedicated for health care centres. The paper presents an intelligent information system of electronic medical records, taking into account governmental plans for implementing the Health Insurance Card and pl.ID systems, as well as e-Prescription system being implemented by the Centre for Health Information Systems. The assumptions of the strategy of implementing IT systems in health protection in Poland focus on achieving a satisfactory level of informatization in the basic areas of the health protection system. So far, Poland has succeeded the first step in implementing a single system, but the next step should be focused on data redundancies and should be coherent, both for patients and health care centres [Direction of inf., 2009].

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

[Apteka euromedica, 2014] http://www.aptekaeuromedica.pl, (accessed 09 April 2014).

[Apteka, 2012] Apteka Euromedica, http://www.aptekaeuromedica.pl, (accessed 09 Januar 2012).

[Blumenthal & Glaser, 2007] Blumenthal D, Glaser JP. Information technology comes to medicine. N Engl J Med.2007, http://www.ncbi.nlm.nih.gov/pubmed/17568035, (accessed 18 April 2014).

[Direction of inf., 2009] Directions of informatisation „e-Health Poland” for the years 2011-2015; Study as part of Update of the document „Strategy e-Health 2004-2006”, which is the subject of the Agreement as part of the project no 2006/018-180.01.01 co-financed by the European Union from the funds Transition Facility 2006; Beneficiary CSIOZ, Warszawa 2009, [WWW document] www.csioz.gov.pl/publikacja.php, (accessed 09 Januar 2012).

[Direction, 2011] Directions of development of e-services in health protection in West Pomeranian Voivodeship for the years 2011-2020; Szczecin 2011, www.wz.wzp.pl/download/index/biblioteka/7620, (accessed 09 April 2014).

[Do lekarza, 2012] Do lekarza z PESELem, czyli e-WUŚ, http://mac.gov.pl/dzialania/e-administracja/do-lekarza-z-peselem/, (accessed 02 December 2012).

[Friedberg et al, 2013] Friedberg M.W., Chen P. G., Van Busum K. R., Aunon F. M., Pham Ch., Caloyeras J. P., Mattke S., Pitchforth E., Quigley D. D., Brook R. H., Crosson F. J., Tutty M., Factors Affecting Physician Professional Satisfaction and Their Implications for Patient Care, Health Systems, and Health Policy, The RAND Corporation, http://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR439/RAND_RR439.pdf, (accessed 09 April 2014).

[Goldzweig et al, 2009] Goldzweig C.L., Towfigh A., Maglione M., Shekelle P.G., Costs and benefits of health information technology: new trends from the literature, 2009 Mar-Apr., http://www.ncbi.nlm.nih.gov/pubmed/19174390, (accessed 30 April 2014).

[Han & Lopp, 2013] Han H., Lopp L. Writing and reading in the electronic health record: an entirely new world. Medical Education Online, Feb 5, 2013, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3566375/, (accessed 09 April 2014).

[Infoklient bi, 2012] http://www.infoklient.asseco.pl/nowe_technologie/system_analityczny_bi, (accessed 09 April 2014)

[Infoklient systemy administracyjno zarządcze, 2014] http://www.infoklient.asseco.pl/oprogramowanie/infomedica/systemy_administracyjno_zarzadcze, (accessed 09 April 2014)

[Informklient systemy medyczne, 2012] http://www.infoklient.asseco.pl/oprogramowanie/infomedica/systemy_medyczne, (accessed 09 April 2014)

[Kalkulation von Fallkosten, 2002] Kalkulation von Fallkosten. Handbuch zur Kalkulation von Fallkosten. Version 2.0 – 31. Januar 2002, Deutsche Krankenhausgesellschaft, 2002, [WWW document] http://www.aok-gesundheitspartner.de/imperia/md/gpp/bund/krankenhaus/drg_system/kalkulation/kalkulationshandbuch_2_0.pdf, (accessed 09 Januar 2012).

[Kamsoft, 2014] http://www.kamsoft.pl, (accessed 09 April 2014)

[Kaplan & Harris-Salamone, 2009] Kaplan B., Harris-Salamone K. Health IT Success and Failure: Recommendations from Literature and an AMIA Workshop, 2009, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2732244/, 2009, (accessed 30 April 2014).

[Karta, 2013] Karta Ubezpieczenia Zdrowotnego, http://www.nfz-katowice.pl/?k0=02_ubezpieczony&k1=04_karta_ubezpieczenia_zdrowotnego&l=kuz-do-czego-sluzy.html, (accessed 18 Januar 2013).

[Kawiorska, 2004] Kawiorska D. Narodowe Rachunki Zdrowia. Kantor Wydawniczy ZAKAMYCZE, Kraków 2004.

[Keyhani et al, 2008] Keyhani S., Hebert P.L., Ross J.S., Federman A., Zhu C.W., Siu A.L., Electronic health record components and the quality of care; Med Care. 2008 Dec., http://www.ncbi.nlm.nih.gov/pubmed/19300317, (accessed 30 April 2014)

[Kralewski et al, 2008] Kralewski JE, Dowd BE, Cole-Adeniyi T, Gans D, Malakar L, Elson B. Factors influencing physician use of clinical electronic information technologies after adoption by their medical group practices. Health Care Manage Rev., http://www.ncbi.nlm.nih.gov/pubmed/18815501, (accessed 19 April 2014).

[Laudon & Laudon, 2011] Laudon K., Laudon J. Essentials of management information systems. Ninth Edition, Pearson, New Jearsey, 2011.

[Linder et al, 2011] Linder J. A., Rigotti N. A., Schneider L. I., Kelley J. H. K., Brawarsky P., Schnipper J. L., Middleton B., and Haas J. S. Clinician characteristics and use of novel electronic health record functionality in primary care. J Am Med Inform Assoc., 2011 Dec., http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3241168/, (accessed 19 April 2014).

[Linder et al, 2012] Linder J.A., Schnipper J.L., Middleton B. Method of electronic health record documentation and quality of primary care, J Am Med Inform Assoc. 2012 Nov-Dec., http://www.ncbi.nlm.nih.gov/pubmed/22610494, (accessed 19 April 2014)

[Menachemi & Brooks 2006] Menachemi N., Brooks RG. EHR and other IT adoption among physicians: results of a large-scale statewide analysis. J Healthc Inf Manag., http://www.ncbi.nlm.nih.gov/pubmed/16903665, (accessed 19 April 2014).

[Menachemi et al, 2006] Menachemi N., Perkins RM, van Durme DJ, Brooks RG, Examining the adoption of electronic health records and personal digital assistants by family physicians in Florida, Inform Prim Care. 2006, http://www.ncbi.nlm.nih.gov/pubmed/16848961, (accessed 19 April 2014)

[Middleton et al, 2013] Middleton B., Bloomrosen M., Dente M.A., Hashmat B, Koppel R, Overhage J.M., Payne T.H., Rosenbloom S.T., Weaver C., Zhang J.; American Medical Informatics Association, Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA, J Am Med Inform Assoc. 2013, http://www.ncbi.nlm.nih.gov/pubmed/23355463, (accessed 18 April 2014).

[Mmedica, 2012] https://mmedica.asseco.pl/Funkcjonalno%C5%9B%C4%87/tabid/55/Default.aspx, (accessed 09 April 2014)

[National Registers System, 2013] Nowa odsłona projektu pl.ID, 2013, https://msw.gov.pl/pl/aktualnosci/11600,Nowa-odslona-projektu-plID.html, (accessed 18 April 2014).

[NFZ, 2012] Departament Współpracy Międzynarodowej Centrali NFZ, https://www.ekuz.nfz.gov.pl/, (accessed 07 July 2012).

[Pankowska, 2004] Pankowska M., Value-driven management in e-healthcare, Stud Health Tech Informat 105:3–11

[pl.ID, 2014] pl.ID – rozpoczyna się cykl 16 spotkań z przedstawicielami urzędów gminnych,2014, http://obywatel.gov.pl/pl-id-rozpoczyna-sie-cykl-16-spotkan-z-przedstawicielami-urzedow-gminnych/4436#sthash.dT5ff9HZ.dpuf, 2014, (accessed 30 April 2014).

[Pollard et al, 2013] Pollard S.E., Neri P.M., Wilcox A.R., Volk L.A., Williams D.H., Schiff G.D., Ramelson H.Z., Bates D.W. How physicians document outpatient visit notes in an electronic health record, 2013 Jan, http://www.ncbi.nlm.nih.gov/pubmed/22542717, (accessed 30 April 2014).

[Poon et al, 2010] Poon E.G., Wright A., Simon S.R., Jenter C.A., Kaushal R., Volk L.A., Cleary P.D., Singer J.A., Tumolo A.Z., Bates D.W. Relationship between use of electronic health record features and health care quality: results of a statewide survey. Med Care. 2010 Mar, http://www.ncbi.nlm.nih.gov/pubmed/20125047, (accessed 30 April 2014).

[Poznaj, 2012] Poznaj system e-Recepta, http://e-recepta.gov.pl/web/sim/poznaj-system-e-recepta, (accessed 09 April 2012).

[Reply of the Secretary, 2013] Reply of the Secretary of State Zbigniew Podraza in Health Ministry – on behalf of the Minister – to the question no 9353 concerning resigning from introducing a medical services register 02.03.2005, http://orka2.sejm.gov.pl/IZ4.nsf/main/7719F1C7, (accessed 18 January 2013).

[Rouf et al, 2008] Rouf E., Chumley HS, Dobbie AE, Electronic health records in outpatient clinics: perspectives of third year medical students, BMC Med Educ. 2008 Mar 31, http://www.ncbi.nlm.nih.gov/pubmed/18373880, (accessed 18 April 2014).

[Shekelle et al, 2006] Shekelle PG, Morton SC, Keeler EB, Costs and benefits of health information technology. Evid Rep Technol Assess (Full Rep). 2006 Apr. (132):1-71., http://www.ncbi.nlm.nih.gov/pubmed/17627328, (accessed 30 April 2014).

[Sidorko, 2012] Sidorko A. Druk RMUA co miesiąc tylko na wyraźne żądanie pracownika, 2012, [WWW document] http://serwisy.gazetaprawna.pl/praca-i-kariera/artykuly/613096,druk_rmua_co_miesiac_tylko_na_wyrazne_zadanie_pracownika.html, (accessed 18 April 2014).

[Sittig & Singh, 2012] Sittig D.F., Singh, H. Electronic Health Records and National Patient-Safety Goals, 2012, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3690003/#!po=50.0000, (accessed 18 April 2014).

[Sittig et al, 2011] Sittig D.F., Wright A., Meltzer S., Simonaitis L., Evans R.S., Nichol W.P., Ash J.S., Middleton B., Comparison of clinical knowledge management capabilities of commercially-available and leading internally-developed electronic health records, BMC Med Inform Decis Mak. 2011 Feb 17, http://www.ncbi.nlm.nih.gov/pubmed/21329520, (accessed 18 April 2014)

[Sołtysik-Piorunkiewicz, 2012] Sołtysik-Piorunkiewicz. A Implementation of IT system in health protection. In (Knosala R, Ed) Innovations in management and production engineering, pp 568-576, Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole 2012.

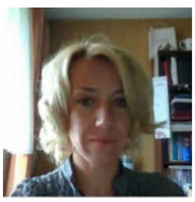
[Strug, 2009] Strug A. Strategia rozwoju systemów informatycznych NFZ. Łódź 2009, http://www.ezdrowie.lodzkie.pl, (accessed 09 April 2014).

[System, 2013] System: Zdrowotny Informator Pacjenta, https://zip.nfz.poznan.pl/ap-zip-user/, (accessed 18 Januar 2013).

[Thatcher, 2013] Thatcher M., IT Governance in Acute Healthcare: A Critical Review of Current Literature. eHealth: Legal, Ethical and Governance Challenges, pp 349-370, http://link.springer.com/chapter/10.1007/978-3-642-22474-4_15, (accessed 18 April 2014).

[Ujejski, 2011] Ujejski M. Funkcje karty ubezpieczenia zdrowotnego realizowane w elektronicznym dowodzie osobistym. In Materiały konferencyjne "IT w służbie zdrowia GigaCon", Warszawa, http://sdcenter.pl/system/lectures/attachment1s/1729/original/Funkcje_KUZ_w_pl.ID_23.03.2011_final_.pdf?1301043192 (accessed 09 Januar 2012).

Authors' Information



Anna Sołtysik-Piorunkiewicz *– Ph.D. is employed at the University of Economics in Katowice as lecturers on Faculty of Informatics and Communication, at Department of Informatics, Katowice, Poland; e-mail: apiorunkiewicz@ue.katowice.pl.*

Major Fields of Scientific Research: computer science, systems analysis and computer system design, management information systems, e-business, e-health and public informatics.

ПАМЯТИ ВИКТОРА ПОЛИКАРПОВИЧА ГЛАДУНА

1936 - 2014

24 апреля 2014 года скончался ***Виктор Поликарпович Гладун*** – доктор технических наук, профессор, старший научный сотрудник Института кибернетики имени В.М. Глушкова Национальной Академии Наук Украины.

Виктор Поликарпович Гладун был ярким представителем талантливой научной молодёжи, которая в начале шестидесятых пришла в кибернетику по зову разума и сердца. С тех пор и до конца дней он оставался верным своему призванию, которое осознавал как научный долг.

Родился Виктор в Киеве в 1936 году в учительской семье. В сорок первом отец ушёл на фронт и погиб под Харьковом. Мама с сыном пережили оккупацию, потом была школа, которую Виктор закончил с золотой медалью.

В 1953 году ***Виктор Гладун*** поступил в Киевский Государственный Университет имени Тараса Шевченко. В этом году был образован новый факультет – радиофизический, преподавание велось на высоком уровне, соответствующем потребностям страны, в которой бурно развивалось ракетостроение и вскоре был сделан первый шаг в Космос. ***Виктор Поликарпович***, благодаря блестящей учёбе, дружелюбию и юмору сразу же завоевал авторитет среди однокурсников, стал признанным комсомольским лидером. Учась в Университете, он завоевал высокие спортивные разряды по гимнастике и альпинизму. Был принят в коммунистическую партию. Ему доверили руководить студенческим отрядом на целине в горячее время уборки урожая летом 1956 года, и этот добровольный трудовой порыв был отмечен рядом благодарностей.

После получения «красного диплома» об окончании Киевского госуниверситета молодой специалист был направлен по распределению на работу в престижный Научно-исследовательский институт радиоэлектроники (г. Киев). Сразу же был замечен и поставлен во главе группы разработчиков системы автоматического управления радиотехническим комплексом. Здесь он научился доводить работы до их приёма Государственной комиссией, приобрёл навыки управления инженерами и техниками в условиях жёсткой производственной дисциплины. Это был его вклад в укрепление обороноспособности страны.

И всё же, несмотря на отличные научные перспективы, ***Виктора Гладуна*** неумолимо тянуло в Институт кибернетики (в 1962 году – Вычислительный Центр Академии Наук УССР), который основал выдающийся учёный и организатор, основатель кибернетики на Украине Виктор Михайлович Глушков. Поэтому после трёхлетнего инженерного стажа в НИИ ***Виктор Поликарпович*** поступил на работу в ВЦ АН УССР – в лабораторию Теории цифровых автоматов, которую возглавлял талантливый учёный, впоследствии доктор технических наук, профессор Зиновий Львович Рабинович. На всю оставшуюся жизнь творческая дружба связала их вместе, в жарких дискуссиях между учителем и учеником рождались идеи, алгоритмы, смелые публикации. Под руководством З.Л. Рабиновича в 1967 году ***Виктор Поликарпович Гладун*** защитил кандидатскую диссертацию «Автоматизированные системы обработки информации».

В атмосфере научного поиска и свободного общения, на бурно протекающих в лаборатории научных семинарах, в которых очень часто принимал участие и директор Вычислительного Центра В.М. Глушков, рождались самые фантастические проекты, которые потом претворялись в жизнь. В.М. Глушков ставил перед молодым коллективом задачу разработки электронно-вычислительных машин для проведения сложных инженерных и теоретических расчётов, которые, в отличие от универсальных ЭВМ, могли бы выпускаться крупносерийно и насытить передовой вычислительной техникой проектно-конструкторские организации, отраслевые институты и конструкторские бюро СССР. В Институте кибернетики были разработаны теоретические и инженерно-конструкторские предпосылки для выполнения этой задачи и практически созданы малые электронно-вычислительные машины. От этой государственной стратегии – создание крупных серий малых ЭВМ – оставался один небольшой шаг до разработки советских персональных компьютеров, но страна не успела сделать этот шаг.

В период с 1962-го по семидесятые годы ***Виктор Поликарпович Гладун*** разрабатывал теорию дискретных автоматов и теорию алгоритмов, включая создание машинных языков высокого уровня, ориентированных на «дружественный» интерфейс пользователя. Параллельно он размышлял о принципах обработки информации машинами «неймановского типа» и сопоставлял их с информационными процессами в биологических системах. Это привело его к мысли, что можно создать алгоритмическую модель, которая в какой-то степени воспроизводила бы процессы преобразования информации в мозге человека. Так родилась фантастическая по научной смелости идея – Растущие Пирамидальные Сети (РПС). Спустя годы мы увидели, что это был прорыв в области создания кибернетических устройств и Искусственного Интеллекта.

В основу РПС положена архитектура нейроподобной сети. Алгоритм работы такой сети предельно прост и потому чрезвычайно эффективен. Повторяющиеся регулярно сигналы (стимулы) на входе нейронной сети образуют в ней новые структуры с устойчивыми связями. Если некоторая комбинация входных сигналов повторяется, то в сети «вырастает» отдельный «нейрон» (узел с входящими и выходящими связями), который эту комбинацию запоминает и впоследствии на неё реагирует. Нейронные структуры нижних уровней РПС постепенно – под воздействием обучающей выборки – объединяются в более сложные структуры верхнего уровня. Таким образом, создаётся иерархическая растущая и развивающаяся организация памяти. После конца обучения образуется сетевая структура, сохраняющая в себе «образ» множества входных сигналов. Этот образ может быть представлен в виде множества логических функций, которые формально отображают закономерности порождения множества сигналов.

Как и многие новые и революционные идеи, Растущие Пирамидальные Сети поначалу были встречены научным сообществом с известным скептицизмом. Это, в частности, проявилось в том, что ***Виктору Поликарповичу*** пришлось защищать свою докторскую диссертацию дважды. Помогли в защите такие качества молодого учёного как упорство в достижении цели и уверенность в созданном им научном продукте. Защита диссертации «Теоретические основы процессов формирования понятий и планирования действий в автоматизированных системах научных исследований», которая состоялась в Киеве, в Институте кибернетики им. В.М. Глушкова в 1983 году, прошла триумфально.

Правильность своей научной теории ***Виктор Поликарпович Гладун*** доказал не только теоретически, но и на практике. Под его руководством коллектив талантливых исследователей в течение 1970 – 2007 гг. разработал ряд инструментальных программных комплексов: КОДЭКС (для построения экспертных систем), КОНФОР (для обнаружения и анализа закономерностей, на основе индуктивного обобщения данных об известных объектах), АНАЛОГИЯ (для решения задачи классификации на основе вывода по аналогии для объектов, имеющих внутреннюю структуру) и других. КОНФОР способен решать задачи обработки данных, в которых значения целевого свойства зависят от многих параметров без априорных предположений о характере такой зависимости. Наряду с этими комплексами разработан уникальный Язык представления знаний – ЭКСПРЕСС, в котором предусмотрены средства для записи формулировок задач, правила преобразования ситуаций, формы записи планов решений и другие макропроцедуры высокого уровня.

Эффективность новых компьютерных технологий была апробирована в таких областях прикладных исследований, как медицинская диагностика, техническая диагностика радиоэлектронных устройств, определение жизненного цикла механических узлов и некоторых других. С помощью программных комплексов на основе РПС решались разноплановые задачи: прогнозирование мест залегания полезных ископаемых, предсказание солнечной активности, задачи выделения закономерностей, которые присущи экономическим ситуациям и технологическим процессам, и другие. Более трех десятков лет (с 1976 г.) система АНАЛИЗАТОР и ее развитие – КОНФОР используются в исследованиях Институте металлургии им. А.А.Байкова (г. Москва) для прогнозирования существования новых материалов с заданными свойствами; кроме этого система показала свою эффективность при решении задачи планирования последовательности действий для синтеза химических соединений (Ин-т органического синтеза, г. Рига). Системы планирования решений – APROS, CODEX применялись в экспериментальном робототехническом комплексе для планирования действий роботов. Прямой народнохозяйственный эффект был получен от внедрения метода прогнозирования миграции рыбы в Баренцевом море (НПО «Севрыба», г. Мурманск, Россия).

В последующие годы деятельности внимание учёного, обладавшего широким кругозором и научным предвидением, сосредоточилось на практических задачах, решаемых человеком на интеллектуальном уровне. Им был сделан шаг от процессов планирования и принятия решений к проблемам понимания языковых сообщений, которые представлены в виде текстов или дискурсов сложной семантической структуры.

Одним из первых был построен программный комплекс КОНСПЕКТ (2000 – 2007 гг.). Его интеллектуальная работа заключается в том, чтобы для заданного текста произвольной тематики составить автоматически конспект, то есть – дать краткий и грамотный пересказ содержания текста с минимальными искажениями его связности. Использование данной информационной технологии даёт пользователю возможность не только хранить массивы документов в сжатом виде (с сохранением их семантики и основного содержания), но и быстро отбраковывать ненужные, выбирать актуальные файлы.

В последние годы жизни ***Виктор Поликарпович Гладун*** поставил перед собой амбициозную научную задачу – построить общение человека и компьютера на принципах взаимопонимания. С самого начала он установил, что такое общение невозможно без представления и оформления системы знаний – онтологии. Развитие этих идей привело к созданию концептуального алгоритма, в базу знаний которого заложены, с одной стороны, знания человека об окружающем его Мире и знания о предметных областях деятельности, а с другой стороны – процедуры синтаксического, семантического и прагматического анализа «живого» текста. Предполагается, что понимание текста достаточно сложной когнитивной структуры достигается через извлечение из него базовой категории сознания и коммуникации – «Смысла». Трудность состояла именно в формализации самого понятия «Смысл». Работы в этом направлении продолжаются. В случае успеха компьютеры следующих поколений, андроиды и аватары будут обладать элементами антропоморфного понимания и мышления.

Виктор Поликарпович Гладун проявил себя также и как хороший организатор науки. Много лет он, будучи доктором технических наук и профессором, возглавлял лаборатории исследователей в НПО «Горсистемотехника» и в Институте кибернетики имени В.М. Глушкова НАН Украины (г. Киев). Особенный период его жизни связан с началом девяностых годов прошлого века, когда в результате распада СССР региональная наука оказалась на грани деградации. В обстановке отсутствия финансирования он не только нашёл в себе силы сохранить лабораторию, продолжать научную работу, но и сумел объединить учёных Украины, России, Беларуси, Болгарии, Польши, Великобритании, США и других стран в единую организацию, которая действует как «Ассоциация создателей и пользователей систем искусственного интеллекта» (АСПИС). Международная конференция, регулярные заседания которой проводятся под названием «Знания – Диалог – Решение» (англ. «KDS» – Knowledge-Dialogue-Solution), известна в научном мире на разных континентах; её значимые труды публикуются на страницах журнала "International Journal on Information Technologies and Applications" ("IJ ITA").

В период 1987 – 1991 гг. ***Виктор Поликарпович*** был председателем международной рабочей группы WG26, а в 1988 – 1992 гг. – сопредседателем международной рабочей группы WG25, являлся членом Совета ассоциации по проблеме «Искусственный интеллект», принимал участие в организации многих научных конференций как руководитель программного комитета или в составе организационных комитетов, входил в редакционные советы международных научных изданий (в том числе – журнала IJ ITA), под его научным руководством развивалось Международное научное сообщество ITHEA

(ITHEA International Scientific Society) и издательство ITHEA (www.ithea.org), которые приобрели мировой известностью и значимостью.

Он вёл преподавательскую работу в ведущих высших учебных заведениях Украины – в Киевском Национальном Университете им. Т.Г. Шевченко, а также в Национальном Университете «НТТУ КПИ». Под его руководством защитились девять соискателей учёной степени кандидата технических наук.

Высокий уровень культуры ***Виктора Поликарповича***, знание языка и общественное доверие позволили ему в шестидесятых и семидесятых годах пройти научную стажировку в Англии и в Италии, сотрудничать с учёными США. Именно поэтому он легко находил общий язык с зарубежными исследователями, которые всегда высоко ценили его эрудицию, интеллигентность, уважение к партнёру. Результаты его научных исследований отражены в пяти монографиях и многочисленных статьях.

Увлечённость наукой и высокое трудолюбие ***Виктора Поликарповича*** гармонично сочетались в его характере с разносторонними увлечениями историей и культурой своей Родины, музыкой, путешествиями. В компаниях друзей он излучал энергию и самобытный юмор, оставаясь в то же время «в тени». Запомнился нам исключительным и чистым стремлением придти на помощь в трудную минуту, поделиться в походе последним куском хлеба. Был прекрасным семьянином, отцом двух сыновей, воспитывал двух внуков.

В наших сердцах останется глубокое сожаление о потере Учителя и Друга, и в то же время –светлое чувство причастности к его жизни. Сохраним тепло дружеского рукопожатия ***Виктора Поликарповича Гладуна*.**

Л.А. Святогор, В.Ю. Величко

Институт кибернетики им. В.М. Глушкова НАН Украины, г. Киев

Список монографий В.П. Гладуна

Глушков В.М., Гладун В.П., Погребинский С.В. Обработка информационных массивов в автоматизированных системах управления. Наукова думка, К, 1970.

Гладун В.П. Эвристический поиск в сложных средах. Наукова думка, К, 1977.

Гладун В.П. Планирование решений. Наукова думка, К, 1987.

Гладун В.П. Процессы формирования новых знаний. София, СД "Педагогь", 1994.

Партнерство с компьютером К.:"Port-Royal", 2000. - 128 с.

TABLE OF CONTENTS