КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ имени АЛЬ-ФАРАБИ

ИНСТИТУТ ВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ
СИБИРСКОГО ОТДЕЛЕНИЯ РАН

ISSN 1560-7534
ISSN 1563-0285

СОВМЕСТНЫЙ ВЫПУСК

по материалам международной научной конференции
"Вычислительные и информационные технологии в науке, технике и образовании"
(CITech-2015)
(24-27 сентября 2015 года)

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ

Том 20

ВЕСТНИК КАЗНУ им. АЛЬ-ФАРАБИ

Серия математика, механика и информатика № 3 (86)

ЧАСТЬ IV

АЛМАТЫ – НОВОСИБИРСК, 2015

Вычислительные Технологии

2015

Том 20

Computational Technologies

2015
Vol 20

ВЕСТНИК КАЗНУ 2015

Серия математика, механика и информатика № 3 (86)

Научное издание

Вестник КазНУ
Серия математика, механика, информатика
№ 3(86) 2015

ИБ № 8517

Подписано 14.04.2015 г. Формат 60x84 1/8. Бумага офсетная.
Печать цифровая. Объем 30,5 п.л. Тираж 500 экз. Заказ № 2620.
Казахского национального университета им. аль-Фараби.
050040, г. Алматы, пр. аль-Фараби, 71, КазНУ.
Отпечатано в типографии издательского дома «Қазақ университеті».
Издательский дом «Қазақ университеті»

The BULLETIN of KAZNU 2015

Mathematics, Mechanics and Informatics Issue **№ 3 (86)**

Table of Contents

Session IV. New Information Technologies in Education

Session V. Technological Process Automation and Control

Session IV. New Information Technologies in Education

Concepts of Activization of Trainees Within Structural Model of Education

D.N. Ashurova, M.U. Raimova, Z.Kh. Yuldashev, and M.A. Yuldasheva

National University of Uzbekistan,
Vuzgorodok, 100057 Tashkent, Uzbekistan
ziyaut@mail.ru
http://www.nuu.uz

Abstract. In this work some approaches to theoretical judgment of problems of education are offered, recommendations about a practical solution of the problem of activization of trainees and, as a result, to increase of level of a reefficiency trained in a context of skill of generation of new knowledge are made.

Validity new option of a paradigm of education: education through all life

Below we will assume acquaintance to work [1], and to use the definitions entered in it and concepts. We will note that in it the new option of a paradigm of education is offered: **education through all life**, and it has rather high validity. As not everyone, probably very laconic phrase from the point of view of a certain circle of people, can be an equivalent of the points of view recognized by experts on problems of a tendency and the system purpose, justification of a validity offered option of a paradigm it is necessary in order that the fresh wording could be perceived further as definition option.

First of all, this option of a paradigm of education, in our opinion, as a dominant on which it is necessary to focus an education system, allocates a tendency of exponential accumulation of knowledge in modern society. This tendency obviously being, a positive consequence of paradox of knowledge ("the more I learn new, the I learn more how a little I know"), supported not only scientific, and sometimes mere curiosity, bears the most valuable fruits not only at the correct organization of science, but also education process. From here the predictive validity of **a paradigm is obvious**.

Objective process of accumulation of knowledge, also objectively generated various tendencies in accumulation of types of knowledge and concrete education systems. "If Ancient Greece was famous for development of that we now call fundamental sciences, Ancient Rome gives extensive literature generally applied character ..." - this rather exact characteristic of the Hellenic and Roman protoscience reflects historical tendencies in accumulation of knowledge. Certainly, a source of any knowledge is the nature, the world surrounding us the phenomena occurring in it and processes. However, natural is as well a question what is the reason of accumulation of knowledge of a certain type which with need are acquired through various education systems? The answer to this question, certainly, multidimensional, but here in connection with an objective we will stop only on some of these reasons.

Despite preferable financing of applied sciences observed presently from state and business of the structures, fundamental science, is more exact its theories, results and conclusions which steadily enrich the maintenance of a scientific picture of the world, objectively remains in the attention center. In the conditions of globalization realization of scientific ideas happens in the shortest terms. It belongs not only to new high, therefore, competitive technologies, but also transnational scientific projects on space research, in the field of ecology, use of renewables,

medicine, secrets of origin of the Universe when efforts of one separately taken country can be insufficient not only from the financial point of view, but also intellectual as the fundamental science, first of all, is international. This positive integration, and also motivational message of paradox of knowledge when the more the person knows, the more it feels a lack of knowledge, along with other factors, promote the mentioned exponential accumulation of knowledge. At all times society was faced by a problem of transfer of the accumulated knowledge or **a problem of training**. Need of the solution of a problem of training was intuitively felt in primitive communities as a security measure for hunger, preservation and sort enhancement when the young generation was trained in hunting, agriculture or permission of other problem situations. The accumulated knowledge of environment, i.e. knowledge and skills of protoscience ancient, was transferred to youth in forms of primitive and evident training and gradually turned into constantly operating schools and systems of educational institutions. In this regard, as the answer to the question posed we come to the known prime cause of any purposeful activity of the person - to need of satisfaction of requirements.

The known phrase, urging to appreciate the book a knowledge source, states importance and value historically the most widespread physical source where information which only in the course of training turns into knowledge of the individual is reflected. Below we will try to express our view of concepts of information and knowledge, - however irrespective of where information or knowledge of processes and the phenomena, objective laws of the nature and separate sciences is reflected, their generator, generating substance is the human thought. Being nonspecialists, we intentionally don't concern neither mechanisms of thought process, nor sources of origin of thought as a unique product, noting only grandiose importance of any thought as source in motivation of knowledge and generation of new knowledge. After all new knowledge generates training process.

It is undoubted that modern education systems, also as well as "ancient", have the common and local problems of conceptual and tool character. To conceptual problems, along with development of concrete innovative techniques of training it is necessary to carry, problems of updating and judgment of a paradigm of education, and also a problem of creation of adequate models of the education having the general and/or national character. The range of tool problems of education, undoubtedly, is wider, takes private techniques of teaching, questions of development of normative documents, educational and methodical literature, questions of a hardware, the software in the conditions of electronic training, ? generally, all that provides educational process in the light of use innovative including information technologies. In this sense if to remember some it is indisputable spread doctrines in training, slogans or options of paradigms the educations accompanying training and trained in specified process of evolution of society and the corresponding education system, it is easily possible to be convinced that this paradigm possesses **a discriminant validization**, - it is easy to distinguish it from others, is **constructive and obvious**.

We will notice that separately taken teachers, and, especially trainees, can precisely not know a number of terms of pedagogical science, not know that those principles and concepts on which real educational process and in which they participate is conducted, correspond to some paradigm just as that Molyerovsky hero who was surprised that "appears, I told all life prose". However modern teachers have to be most familiar and with the last achievements of pedagogical science, and with modern pedagogical technologies, and with opportunities of application of information technologies in education.

Definition. Property to open knowledge which are unknown to this personality or are new and unknown to anybody, we will call property of a reefficiency of knowledge.

Below in this work some approaches to theoretical judgment of the mentioned problems of an education system of conceptual and tool character, and also the recommendation about a practical solution of the problem of activization of trainees and, as a result, increase of level of **a reefficiency of knowledge trained** in a context of skill of generation of new knowledge are offered.

The problems of an education system connected with its informatization

In the course of training each pupil systematically receives information. One of basic problems of an education system is the problem of transformation *of perceived information in knowledge, skills*. It **is obvious** that knowledge has abstract character, is expressed in concepts and categories of the corresponding branch of science, traditionally fixed on data carriers while carriers of skills are people. Memory of the person also fixes knowledge, but has property to forget them. Fixing of knowledge in long-term memory is promoted by repeated reference to concrete knowledge. In turn memory of the person, participating in thought processes, creates a basis for the solution of pressing problems and realization of unique ability of the person on generation of new knowledge. Thus, **knowledge in pure not modelled option is inherent in only the reasonable person** [2] as it exploits them on the basis of intellectual processing. Any knowledge can become information if it is recorded on carriers and isn't used in interests of the person. To the contrary, as soon as information is used for the solution of problems, it turns into knowledge. Such treatment of distinction between information and knowledge equally concerns also cases of modeling of thought processes, in particular information use at the solution of tasks on the COMPUTER. In concept **skill** is put the certain practical experience of the personality gained in the course of the repeated solution of the same task. **Ability of is ability** creatively to apply skills on the knowledge base. As a rule, the part of knowledge, skills even during life of one person repeatedly becomes outdated and demands updating. The objective process of accumulation of knowledge noted above, is more exact a demand of new knowledge, for an education system sets other task, namely a problem of instilling in each pupil of abilities on updating and continuous accumulation of knowledge while processes of development of new technologies demand acquisition of new skills. These reasons, are more exact than a task, generate the general for all education systems of a problem of conceptual and tool character.

To a number of problems of conceptual character refer commitment to a certain paradigm of education, meaning by it a choice of the corresponding concepts and methodologies, and also activity within a certain model of education. In order to avoid repetitions, we will note that by authors in work [2] are stated their point of view and preferences in the matter owing to what further we will concern a question of a choice and implementation of the concept of training.

Despite existence in pedagogics of numerous concepts of training [3-6], first of all it is necessary to allocate those from them which really promote activization of trainees, to origin of abilities in them independently to update available and to get new knowledge. In this regard it should be noted a special role of the main person involved in any education system - *the teacher* who in [2] is defined as *"studying preparing, organizing and carrying-out educational process"*. After all even the young parent who is sincerely wishing not only to grow up the child, but learns to bring up it to how to make it. When the speech comes about the teacher, whose professional duties include formation of knowledge and skills irrespective of where educational process is for what purpose organized, he objectively and regularly is engaged in expansion of baggage of knowledge. Unfortunately, to politicians and the officials, urging to appreciate work of the teacher, not always it is possible to be consecutive in realization of these quite often sounding appeals. Very often there are numerous options of excuses and "the objective reasons". However

here it would be desirable to state, rather simple truth: **pupils can't appreciate knowledge if society doesn't appreciate bearing this knowledge**. Nevertheless, the teacher to be demanded, steadily combines a role of the teacher with a role of the pupil that confirms the statement about recurrence, isolation *on itself and openness to new knowledge* of any educational process. If the question of increase of professionalism or qualification of the teacher traditionally was solved through obligatory courses special, as a rule, now statements and appeals about need of increase of a role and quality of system of traditional system of professional development even more often sound. As an obligatory component of educational process of this system instilling in teachers of skills of use of information technologies *is specified by preparation, the organization and carrying out educational process*. It is undoubted that the question of increase of a role and quality of system of traditional system of professional development belongs to problems of education actual at all times as a whole which can be considered recently in the light of a computerization of an education system, application in educational process of information technologies. More and more persistent there are appeals to expansion of remote education, and here and there and to total transition to such form of education. Discussions in this question inflame and in the course of discussions opposite statements and appeals quite often express, beginning from full refusal and unacceptability of application of information technologies in education, before total and obligatory transition to electronic training [7,8]. It is obvious that the truth, as always, lies somewhere in the middle and consists in rational use of information technologies in an education system, in directive introduction of techniques which have only justified and forms of education.

Coming back to a question of a choice of this or that concept or a training technique, we will note that it depends both on a branch of science, and on the subject, and on the average level of concrete academic group. As practice is criterion of truth, and practice, is more exact life, in its problem situations and tasks very and is very various, a choice of the concrete concept of a technique "once and for all", from our point of view, is hardly possible. Certainly, each teacher needs to know how it is necessary to give lectures and to carry out a practical work, to organize independent work of pupils, in particular students. Than more teacher is informed on the last achievements of pedagogical science even if he is very fruitful scientist, than more he is familiar with experience of teachers of innovators, the more deeply and each occupation irrespective of a form and a look is creatively worked, the he manages to impart more knowledge, the more skills manage to impart the trainee.

"Technique of basic tasks" as way of permanent concentration of attention and activization of trainees

In literature a number of effective techniques on activization of trainees, instilling in them of strong knowledge and the skills based both on internal communications of concrete discipline, and on external relations of studied problems, characteristic subject domain of studied science [9] is known. However, at all diversity, the specified techniques were developed for high school, and approved, is more exact "received a start in life" through approbation on the school contingent and school subjects. At the higher school, except for separate examples of the uncommon teachers scientists connected with creativity, quite often natural it is considered a campaign in an assessment of suitability of the teacher on the basis of usually unilateral criterion: *he is the recognized scientist and the expert in the field, "the good companion", and students at desire will understand it*. Thereby the axiom accepts: "application of a concrete technique and pedagogical technologies for teachers of the higher school business voluntary". Obviously, similar approach not only forms the point of view about minority of a technique in system of the higher education

thereby blocks the road to the pedagogical initiatives, a special technique, concrete ways of activization of trainees. We will give an example of one approach called us **"a technique of basic tasks"** [10].

Subject: Informatics.

Sequence of tasks:

1) To write algebraically admissible expression including 4 constants, two simple variables, having designated them through x and at, two variables with the indexes, three appeals to standard functions.
2) To write this expression by means of algorithmic language (Pascal, C ++, etc.).
3) About to tabulate the specified function.
4) To write procedure of tabulation of primary expression.
5) To write procedure function defining a maximum of elements of the specified table.
6) To write an interpolation polynom according to the made table and in several points to carry out interpolating and extrapolation.
7) To make the program on application of a method of Euler and Runge-Kutt's method, having accepted the made expression as the right part in Cauchy's corresponding task.
8) To initialize a square matrix, having accepted as its elements of value of the made expression having put $x_i = i * h$, $y_j = j * k$ $(i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n)$.
9) For the initialized matrix to solve a problem of own values.
10) On the basis of the initialized matrix to make system of the linear algebraic equations, having taken as the right part one of its columns, to solve the made system methods.

Similar didactic recommendations are developed and approved for the university course "Algebra and Theory of Numbers", the subject Mathematics (the 7th class, the program for specialized training at auxiliary school) [11]. As a whole the specified technique can be considered and as a private technique. For example, application of this technique in linguistics, in particular when studying bases of some living language, **a technique of basic tasks** can be expressed by the following chain of tasks: *to think up a predicate → to think up a subject → to think up two or three adjectives → to think up a verbal adverb phrase → to make a complex sentence.* Thus each pupil thinks out own options, they are discussed, at a loss pupils receive the help.

References

1. Yuldashev, Z.Kh., Ashurova, D.N.: Innovative-Didactic Program Complex and New Formalized Model of Education. Journal of Mathematical Sciences 6(1): 97-103 (2012)
2. Ashurova, D.N., Yuldashev, Z.Kh.: Problems of training of IT specialists in Uzbekistan. "Science and education informatization No. 2(14), (2012)
3. Selevko, G.K.: Modern educational technologies. - M.: National education, -256 pages, (1998)
4. Bespalko, V.P.: Pedagogics and progressive technologies of training. - Moscow: IRPO, -336 pages, (1996)
5. Makhmutov, M. I.: Theory and practice of problem training. Kazan, (1972)
6. Creative pedagogics: methodology, the theory, practice / Under the editorship of Kruglov Yu.G. - M.: MGOPU of M. A. Sholokhov, the Alpha, -240 pages, (2002)
7. Starodubtsev, V.A.: Computer and multimedia technologies in natural-science education. Publishing house HANG-GLIDER, Tomsk: -223 p. (2002)
8. Yuldashev, Z.Kh., Ashurova, D.N.: Application of innovative and information technologies in an education system as time requirement. National education journal, Tashkent, No. 1, page 15-19 (in the Uzbek language), (2006)
9. Ashurova, D.N., Bokiyev, R.R., Yuldashev, Z.Kh.: Media education and problems of increase of activity of trainees//Pedagogical skill. - Bukhara, - No. 2 (in the Uzbek language), (2009)
10. Shatalov, V.F.: Point of support. - M.: Pedagogics, -160 pages, (1987)
11. Yuldashev, Z.Kh., Yuldasheva, M.A.: Mathematics and information technologies a tool for rehabilitation of disabled people. Book of Abstracts IV CONGRESS OF THE TURCKIC WORLD MATHEMATICAL SOCIETY, Baku, p.506, (2011)

Kazakh Morphological Analysis for Statistical Machine Translation: A Case Study

Amandyk Kartbayev and Yeldar Bekbolatov

Laboratory of Intelligent Information Systems, Al-Farabi Kazakh National University, Almaty, Kazakhstan
{a.kartbayev,yeldar.bekbolatov}@gmail.com

Abstract. In this paper, we present results from a set of experiments to determine the effect on translation quality depends on the particular kind of morphological preprocessing that can be represented by finite-state transducers. A high agglutinative nature of the Kazakh language under the condition of poor language resources makes an issue in the derivational morphology processing. Our methods focus on useful phrase pairs in word alignment and provide a mostly language independent approach, which may improve translation into other morphologically complex languages. We processed our algorithms over the Kazakh Wikipedia dump of about 1,5 million unique lexemes and 230 million words overall. Our best translation system increases 3 BLEU points over the Kazakh-English baseline on a blind test set.

Keywords: word alignment, morphology, kazakh word segmentation, machine translation.

Introduction

Morphological segmentation process intended to break words into morphemes, which are the basic semantic units and a key component for natural language processing systems. This is our current subtask in the machine translation project and we also desired to show that a simple segmentation scheme can perform pretty well as the most sophisticated one. Most papers in statistical machine translation(SMT) oriented morphology analysis presents experiments that they consist of numerous experimentation to choose the best among a set of segmentation schemes. These morphological preprocessing schemes focused on various level of decomposition and compare the resulting translation performances, but usually use a subset of morphology and apply only a few simple rules in a segmentation process. In the paper, well known to the intended audience, El-Kahlout and Oflazer[1] explored this task for English to Turkish translation, which is an agglutinative language as Kazakh. Their methods used in the survey were a morphological analyzer and token disambiguation, though translation models trained throw morphemes obviously degrades the translation quality. But they outperformed the baseline results after some morpheme grouping techniques. A research more relevant to this work was done by Bisazza and Federico[2]. Our segmentation model incorporates simple ideas inspired by finite state features such as morphemes and their contexts that range of situations lexeme is likely a morpheme, as any other cases it is a word boundary. We develop a segmentation schemes using syntactic and morphological rules are implemented as finite-state transducers. We focus on derivational morphology and tested our approach on Kazakh wiki and news datasets, which was crawled from Web. The affix system, which will be the focus in this paper, is described in more detail in Table 1.

Our system, using monolingual features only, is one of the most realistic application for Kazakh and compared to Morfessor tool[3], so it can be readily applied to supervised and semi-supervised learning of morphological inflations of language even on speech processing. Also morphological adjustment gives a improved statistical machine translation performance over the pair of the morphological rich and poor languages. A substantial improvement in translation

Table 1. An example of Kazakh agglutination

Stem	Plural affixes	Possesive affixes	Case affixes
stem[kol'+]	plural[+der]	1-st pl.[+imiz]	locative[+de]
stem[kol'+]	-	1-st s.[+im]	locative[+de]
stem[kol'+]	-	-	locative[+de]

performance is achieved, when we used word alignments learned from the data after output of the processing technique, but we found that some of the segmentation errors are caused by morphological analyzer. These kind of errors could be avoid using the data selection, which demonstrates the ability of the method fix it successfully. Using a morphology analysis we out grammatical features of word and can find syntactic structure of input sentence, which further demonstrates the benefit of using this method in machine translation. In this paper, we present a systematic comparison of preprocessing techniques for a Kazakh-English languages pair. Previous researches that we explored on our approaches are rule-based morphological analyzers[4] that consist in deep language expertise and a exhaustive process in system development. Unsupervised approaches use actually unlimited supplies of text to cover very few labeled resource and it has been widely studied for a number of languages[5]. However, existing systems are too complicated to extend them with random overlapping dependencies that are crucial to segmentation. On our general task we refer to the methodology exposed by Oflazer and El-Kahlout on the Turkish-English task. Because, Turkish is also morphologically rich language like Kazakh and not all affix combinations looks grammatical. This means that linguistic knowledge is the key to finding significant segmentation schemes among many possible combinations of the rules. Only rule-based approaches are provided and have done detailed analyses of the Kazakh morphological parsing task. For a comprehensive survey of the rule-based morphological analyze we refer the reader to the research by Altenbek[6] and Kairakbay[7]. The paper is structured as follows: Section 2 discusses the key challenges of translating Kazakh to English. In Section 3 we described the different segmentation techniques we study. And Section 4 presents our evaluation results.

Translation task

In our work we experiment with a range of segmentation technique totally giving a five best distinct schemes. Our results show that the proper selection of the segmentation scheme has a significant impact on the performance of a phrase-based system in a large corpora. The translation experiments described in this paper are carried out with a standard phrase-based Moses[8] system (not Experiment Management System) and the target-side language models were trained on the MultiUN[9]corpora. Generally, breaking up the process of generating the data into smaller steps, modeling the smaller steps with probability distributions, and combining the steps into a coherent story - is called generative modeling. The phrase-based models are generative models that translate sequences of words in "f"into sequences of words in "e"(1), in difference from the word-based models that translate single words in isolation.

$$P(e \mid f) = \sum_{a}^{j} P(e, a \mid f) \tag{1}$$

Improving translation performance directly would require training the system and decoding each segmentation hypothesis, which is computationally impracticable. That we made various kind of

conditional assumptions using a generative model and decomposed the posterior probability(2). In this notation "e"and "f"point out the two parts of a parallel corpus and "a"marked as the alignment hypothesized for "f".

$$P\left(e_1^J, a_1^J \mid f_1^I\right) = \frac{f}{(I+1)^J} \prod_{j=1}^{J} p\left(e_j \mid f_{aj}\right) \tag{2}$$

The use of phrases as translation units is motivated by the observation that sometimes one word in a source language translates into multiple words. Because, a Kazakh word can correspond to a single English word, up to phrases of various lengths, or even to a whole sentence as shown in Table 2. Our objective is to produce alignments which can be used to build high quality machine translation systems[10]. These are pretty close to human annotated alignments that often contain m-to-n alignments, where several source words are aligned to several target words and the resulting unit can not be further decomposed. Using segmentation, we describe a new generative model which directly models m-to-n non-consecutive word alignments. There is a very small improvement in alignment if a source word occurs only once in the parallel text, the probability assigned to it generating each of the words to which sentence is paired will be too high. This problem is solved by smoothing the word-to-word translation probabilities with a coincident distribution.

Improving word alignment

In order to look through this task, we did a series of experiments and found morpheme alignment can be employed to increase the similarity between languages, therefore enhancing the quality of machine translation for Kazakh-English language pair. Our experiments consist of two parts: one is on Kazakh-English morphological segmentation; the other is a case study of the benefits of morpheme based alignment. We use following heuristic methods that improve the generative models for phrase alignment. At first, the tags were assigned to the obtained phrase pair pieces, then we make classification and clustering the phrases according their contexts, also we extract phrase pairs that are not linked within the word alignments, like the phrases containing multiword entities that can not be correctly aligned. We obtained word alignments in both translation directions by the GIZA++ toolkit[11], which is based on the IBM models[12]. We prefer a grow-diag-final symmetrization method to others for both alignment directions. As the first part of our experiments we morphologically segmented Kazakh input sentences to compute morpheme alignment. For these purposes we used Morfessor, an unsupervised analyzer and Helsinki Finite-State Toolkit (HFST)[13]. Helsinki Finite State Toolkit is an open-source implementation of the Xerox finite-state toolkit, that implements the lexc, twol and xfst formalisms for modeling morphlogical rules. After these Kazakh stems and suffixes converted into labeled morphemes, as well as particular English verbs. We append a plus sign to the end of each open tag to know boundaries of internal morphemes from final ones, e.g., [stem+] and [stem] are assumed as different tokens.

Morphological segmentation

Our preprocessing jobs start from morphological segmentation, which includes running Morfessor tool and HFST to each entry of the corpus dictionary. The first step of word segmentation aims to get suffixes and roots from a vocabulary consisting of 1500k unique word forms taken from Kazakh Wikipedia dump[14]. Accordingly, we take surface forms of the words and generate their

all possible lexical forms. In the Kazakh language, as in other agglutinative languages, morphemes are affixed to the root due to the morphotactic rules of the language. These morphotactic rules define the states and the suffixes that can be added to a stem, then change the state of the affixed word. These rules often represented by a certain finite state transducers. Where the transitions are marked as the derivational morphemes, that come in same order as the affixation of the word. Also we use lexicon to label the initial states as the root words by parts of speech such as noun, verb, etc. The final states represent lexeme created by affixing morphemes in each further states. The schemes presented below are different combinations of outputs determining the removal of affixes from the analyzed words. The baseline approach is not perfect since a scheme includes several suffixes incorrectly segmented. In this case, we mainly focused on detection a few techniques for the segmentation of such word forms. In order to find an effective rule set we tested several segmentation schemes named S[1..5], some of which have described in the following Table 2.

Table 2. The segmentation schemes

Id	Schema	Examples	Translation
S1	stem	el	state
S2	stem+case	el + ge	state + dative
S3	stem+num+case	el + der + den	state + num + ablativ
S4	stem+poss+	el + in	state + poss2sing
S5	stem+poss+case	el + i +ne	state + poss3sing + dative

Nominal cases that are expected to have an English counterpart are split off from words: these are namely dative, ablative, locative and instrumental, often aligning with the English prepositions to from, in and with/by. The remaining case affixes nominative, accusative and genitive are not have English counterparts. After treating case affixes we split of possessive suffixes from nouns of all persons except the 1st singular, which doesn't need removed. There are large amount of verbs presenting ambiguity during segmentation, as suppositional verbs 'eken' - 'to seem' and 'goi'. Which do not take personal endings, but follow conjugated main verbs. The verb 'to become' has the forms 'bolu' - 'to become', 'bolar' - 'will become', and 'bolmau' - 'to not become'. There are also the verbs 'bar' - 'to exist/have' and 'jok' - 'to not exist/not have'. These are special verbs because they do not take personal endings. Also a verbs generally refer to group action, e.g. 'oinasu' - 'to play together', 'soilesu' - 'to converse' produce an ambiguity, e.g. a stem 'soile' - 'say' and a suffix 'su' - 'water'. During the process, we hardly determined the border between stems and inflectional affixes, especially when the word and the suffix matches entire word in the language. For instance, a progressive auxiliary word 'jat' - 'alien' and the negation morphemes like 'ba', 'ma', etc, though an irregular form of several verbs. Under many situations, the type of words, which we described, made an inaccurate stemming. In fact, there are lack of syntactic information we cannot easily distinguish among similar cases. While GIZA++ tool produces a competitive alignment between words, the Kazakh sentences must be segmented as we already have in the first step. Therefore our method looks like an word sequence labeling problem, the contexts can be presented as POS tags for the word pairs.

Alignment Model

We extend the alignment modeling process of Brown et al. at the following way. We assume the alignment of the target sentence 'e' to the source sentence 'f' is 'a'. Let 'c' be the tag(from

Penn Treebank) of 'f' for segmented morphemes. This tag is an information about the word and represents lexeme after a segmentation process. This assumption is used to link the multiple tag sequences as hidden processes, that a tagger generates a context sequence 'c' for a word sequence 'f'(3).

$$P\left(e_1^I, a_1^I \mid f_1^J\right) = P\left(e_1^I, a_1^I \mid c_1^J, f_1^J\right) \tag{3}$$

Then we can show Model 1 as(4):

$$P\left(e_1^I, a_1^I \mid f_1^J, c_1^J\right) = \frac{1}{(J+1)^I} \prod_{i=1}^{I} p\left(e_i \mid f_{a_i}, c_{a_i}\right) \tag{4}$$

The training is carried out in the tagged Kazakh side and the untagged English side of the parallel text. If we estimate translation probabilities for every possible context of a source word, it will lead to problems with data sparsity and rapid growth of the translation table. We applied expectation maximization(EM) algorithm to cluster a context of the source sentence using similar probability distributions, avoiding problems with data sparsity and a size of the translation table another case.

Table 3. Part of Speech tag patterns

Tag	Sample	Tag	Sample
NN (Noun)	"el"-"state"	JJS (Adjective, super.)	"tym"-"most"
NNP (Proper noun)	"biz"-"we"	VB (Verb, base form)	"bar"-"go"
JJ (Adjective)	"jasyl"-"green"	VBD (Verb, past tense)	"bardy"-"went"
JJR (Adj, comp.)	"ulkenirek"-"bigger"	VBG (Verb, gerund)	"baru"-"to go"
RB (Adverb)	"jildam"-"speedy"	CC (Conjunction)	"jane"-"and"

We estimate the phrase pairs that are consistent with the word alignments, and then assign probabilities to the obtained phrase pairs. Context information is incorporated by the use of part-of-speech tags in both languages of the parallel text, and the EM algorithm is used to improve estimation of word-to-word translation probabilities. After we use association measures to filter infrequently occurring phrase pairs by log likelihood ratio 'r' estimation[15]. For 'n' pairs of the phrases, we can obtain the phrase pairs whose comparative values are larger than a threshold value as follows(5):

$$R\left(f, e\right) = \frac{r\left(f, e\right)}{Max_e r\left(f, e\right)} \tag{5}$$

Our algorithm, like a middle tier component, processes the input alignment files in a single pass. Current implementation reuses the code from `https://github.com/akartbayev/clir` that conducts the extraction of phrase pairs and filters out low frequency items. After the processing all valid phrases will be stored in a phrase table and be passed further. This algorithm proposes refines by adding morphological constraints between the direct and reverse directions of the alignment, which may improve the final word alignments.

Evaluation

Though our final objective is an improvement of the translation quality of a SMT system, we evaluate the alignment relies with the phrase-based system on the Kazakh-English parallel

corpus of approximately 60K sentences, which have a maximum of 100 morphemes. Our corpora consists of legal documents from `http://adilet.zan.kz`, a content of `http://akorda.kz`, and Multilingual Bible texts. We conduct all experiments on a single PC, which runs the 64-bit version of Ubuntu 14.10 server edition on a 4Core Intel i7 processor with 32 GB of RAM in total. All the experiment files were processed on a locally mounted hard disk. Also we expect the more significant benefits from a larger training corpora, therefore we are in the process of its construction.

We did not have a gold standard for phrase alignments, so we had to refine the obtained phrase alignments to word alignments in order to compare them with our word alignment techniques.

Table 4. Alignment quality results

System	Precision	Recall	F-score	AER
Baseline	57.18	28.35	38.32	36.22
Morfessor	71.12	28.31	42.49	20.19
Rule-based	89.62	29.64	45.58	09.17

Table 4 shows the change in alignment error rate (AER) of the alignments, that the improved model produce a decrease in AER and leads to a better translation quality, measured by BLEU score[16]. A high recall apparently improves translation quality, but low precision may decrease it and a relation between recall and precision is substantial. A high recall and low precision in alignment pretty significant for the amount of generated phrases. The best situation takes place on well maintained recall and precision, which is a result of our study.

Table 5. Metric scores for all systems

System	BLEU	METEOR	TER
Baseline	30.47	47.01	49.88
Morfessor	31.90	47.34	49.37
Rule-based	33.89	49.22	48.04

We employed an approach of the morpheme-based representation as explained in Section 3 about the morphological analysis, which impacts an improvement of +2 BLEU points. The system parameters were optimized with the minimum error rate training (MERT) algorithm [17], and evaluated on the out-of and in-domain test sets. Monolingual corpora from News Commentary was partially used, when we trained 5-gram language models. All language models were trained with the IRSTLM toolkit [18] and then were converted to binary form using KenLM [19]

Table 5 visualizes the best BLEU scores, which were computed using the MultEval [20]: BLEU, TER[21] and METEOR[22]; and we ran Moses three times per experiment setting, and report the highest BLEU scores obtained. Our survey shows that translation quality measured by BLEU metrics is not strictly related with lower AER.

Conclusions

In this work, we address a morpheme alignment problems concerned highly inflected languages. We compared our approach against a baseline of the Moses translation pipeline and another common approach to inflected languages segmentation. By using our method for phrase selection we were able to obtain translation quality better than the baseline method produce, while the phrase table size and the noise phrase pairs have been reduced by substantial level. Although memory requirements of the processing environment are increased, but they are still within manageable limits.

Our method is comparable to other language-specific works, and there are many possible directions for future research. As our approach may produce improvements in alignment quality, any downstream changes of the translation model also possible. We learned that processing the features are integrated into a standard phrase table is an area for improvement. That was our initial investigation into alignment models and further translation experiments will be carried out.

References

1. Oflazer K., El-Kahlout D.: Exploring Different Representational Units in English-to-Turkish Statistical Machine Translation. Proc. of the 2nd Workshop on Statistical Machine Translation, 25-32 (2007)
2. Bisazza A., Federico M.: Morphological Pre-Processing for Turkish to English Statistical Machine Translation. IWSLT 2009, 129-135 (2009)
3. Creutz M., Lagus K.: Unsupervised models for morpheme segmentation and morphology learning. ACM Transactions on Speech and Language Processing, 4(1):Article 3, (2007)
4. Beesley K. R., Karttunen L.: Finite State Morphology, Palo Alto, CA: CSLI Publications, (2003)
5. Goldsmith J.: Unsupervised learning of the morphology of a natural language. Computational Linguistics, 153-98 (2001)
6. Altenbek G., Xiao-long W.: Kazakh Segmentation System of Inflectional Affixes. Proceedings of CIPS-SIGHAN Joint Conference on Chinese Language Processing, 183-190 (2010)
7. Kairakbay B.: A Nominal Paradigm of the Kazakh Language. Proceedings of the 11th International Conference on Finite State Methods and Natural Language Processing, 108-112 (2013)
8. Koehn P., Hoang H., Birch A., Callison-Burch C., Federico M., Bertoldi N., Cowan B., Shen W., Moran C., Zens R., Dyer C., Bojar O., Constantin A., Herbst E.: Moses: Open source toolkit for statistical machine translation, in Proc. of the 45th Annual Meeting of the Association for Computational Linguistics, 177-180 (2007)
9. Tapias D., Rosner M., Piperidis S, Odjik J., Mariani J, Maegaard B., Choukri Kh., Calzolari N.: MultiUN: A Multilingual Corpus from United Nation Documents, in Proc. of the Seventh conference on International Language Resources and Evaluation, La Valletta, Malta, 868-872 (2010)
10. Moore R.: Improving IBM word alignment model 1, In ACL, Barcelona, Spain, 518-525 (2004)
11. Och F.J., Ney H.: A Systematic Comparison of Various Statistical Alignment Models, Computational Linguistics, volume 29, number 1, 19-51 (2003)
12. Brown P.F., Della-Pietra V., Del-Pietra S., Mercer R.L.: The mathematics of statistical machine translation: Parameter estimation, Computational Linguistics, vol. 19, 263-311 (1993)
13. Linden K., et al.: HFST - framework for compiling and applying morphology, Systems and Frameworks for Computational Morphology, Springer-Verlag GmbH, 67-85 (2011)
14. Gabrilovich E., Markovitch Sh.: Computing Semantic Relatedness using Wikipedia-based Explicit Semantic Analysis, Proceedings of The 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, 1606-1611 (2007)
15. Dunning T.: Accurate methods for the statistics of surprise and coincidence, Computational Linguistics, 61-64 (1993)
16. Papineni K., Roukos S., Ward T., Zhu W.: BLEU: a method for automatic evaluation of machine translation, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, 311-318 (2002)
17. Och F.J.: Minimum error rate training in statistical machine translation. Proc. of the 41st Annual Meeting of the Association for Computational Linguistics, 160-167 (2003)
18. Federico M., Bertoldi N., Cettolo M.: IRSTLM: an Open Source Toolkit for Handling Large Scale Language Models, Proceedings of Interspeech, Brisbane, Australia, 1618-1621 (2008)

19. Heafield K.: Kenlm: Faster and smaller language model queries. In Proceedings of the Sixth Workshop on Statistical Machine Translation, 187-197 (2011)
20. Clark J., et al.: Better Hypothesis Testing for Statistical Machine Translation: Controlling for Optimizer Instability, Proceedings of the Association for Computational Lingustics, 176-181 (2011)
21. Snover M., Dorr B., Schwartz R., Micciulla L., and Makhoul J.: A Study of Translation Edit Rate with Targeted Human Annotation, Proceedings of Association for Machine Translation in the Americas, 223-231 (2006)
22. Denkowski M., Lavie A.: Meteor 1.3: Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems, Proceedings of the EMNLP 2011 Workshop on Statistical Machine Translation, 85-91 (2011)

A Concept Map Approach to Supporting Adaptive e-Learning

M. Mansurova,[1] A. Nugumanova,[2] Ye. Baiburin,[2] D. Zyryanov[2]

[1] Al-Farabi Kazakh national university, Almaty, Kazakhstan
[2] D. Serikbayev East Kazakhstan State Technical University, Ust-Kamenogorsk, Kazakhstan
mansurova01@mail.ru,yalisha@yandex.kz,ebaiburin@gmail.com
http://www.kaznu.kz

Abstract. Concept map is a powerful tool of learning of reality which is actively used in the last decade in e-learning. It visualizes domain knowledge in the form of a diagram showing the relationship between domain entities, concepts or ideas. Concept maps can be used to represent different sides of the studied object, to reveal inner structure of the object with varying degrees of detail, to select or hide certain aspects. Therefore, concept maps can be customized depending on the needs of the learner, on the study level and on the all that involves individual learning path. Such an interpretation of concept maps involves automatic way of their construction. Concept maps should be built "on the fly in view of what learning textbooks are chosen, what goals are set, what problems will be solved. The purpose of this paper is to describe an approach based on the mechanism of automatic creation of concept maps which are supporting elements in an individual learning path. The most important element of this mechanism is an ability to parameterize requirements put forward to the presentation of the learning material, taking into account student's individuality. The novelty of the work consists in the step by step description of the proposed mechanism that can be used in all e-learning systems involving adaptive learning.

Keywords: Concept maps, e-learning, NLP.

Introduction

Concept map is a term that was first introduced in the works of Joseph Novak [1,2] to refer to the schemes used in the teaching of complex disciplines containing a huge conceptual apparatus. Currently, the concept maps firmly anchored in teaching practice and they widely used in e-learning as a powerful tool for design and visualization of knowledge. Concept maps visualize domain knowledge in the form of a diagram showing the relationship between different domain entities, concepts or ideas. Concept maps can be used to represent different sides of the studied object, to reveal inner structure of the object with varying degrees of detail, to select or hide certain aspects; here everything depends on the point of view of the author of concept map and on a purpose of the concept map. Therefore, concept map is a very flexible learning tool that can be adjusted to the learning objectives, to the current learning stage, to the level of complexity and to the personal needs of the student, i.e. all included in the concept of "individual learning path". The implementation of this dynamic nature of learning involves largely only one way of concept maps constructing – automatic method. Concept maps should be built "on the fly"with the information about individual learning path.

The purpose of this paper is to describe an approach based on the mechanism of automatic creation of concept maps which are supporting elements in an individual learning path. In our opinion, the most important element of this mechanism is its ability to parameterize requirements put forward to the presentation of the learning material, taking into account student's individuality. The novelty of the work consists in the step by step description of the proposed mechanism that can be used in all e-learning systems involving adaptive learning.

The rest part of the paper is organized as follows. Section 2 describes the method of automatic generation of concept maps based on the teaching textbooks. In Section 3, we propose a new

approach to parameterization of generated concept map, based on the student's model (profile). Section 4 presents the architecture of adaptive learning system based on the use of concept maps. Section 5 contains conclusions and plans for future research.

Automatic generation of concept maps on the basis of teaching textbooks

The proposed method of automatic generation of concept maps on the basis of teaching textbooks consists of three main steps. The first step is an automatic extraction of domain concepts (entities). We considered keywords closely related to the domain as such concepts. To find and select those keywords, we can use a variety of statistical criteria detailed in [1, 2, 3].

One of the most popular criteria for keywords selection is Pearson's criterion which tests the independence of two events. With regard to the keyword selection Pearson's test evaluates the independence of a word and domain. The null hypothesis of Pearson's test is that there is no relation between the word and the domain, which means that the word is distributed with similar frequency among domain texts and among common texts. Accordingly, an alternative hypothesis is that there is a stable relation between the word and the domain, which means that the frequency of the word in domain texts is higher than th frequency in the common texts. Thus, the Pearson's test compares the distribution of words in two sets of documents: a positive set (texts of domain textbooks) and negative set (common texts). The formula estimating this distribution for each word is as follows:

$$X^2 = \frac{(A+B+C+D)(AD-BC)^2}{(A+B)(A+C)(B+D)(C+D)} \tag{1}$$

where A - is a quantity of texts of positive set, containing this word, B - is a quantity of texts of negative set, containing this word, C - is a quantity of texts of positive set, not containing this word, D - is a quantity of texts of negative set not containing this word. Pearson's test is a highly effective and rather simple statistical technique to select keywords but it has one drawback: it requires an alternative (negative) set of documents. If there is no alternative set of documents, the common TF-IDF measure is used to select high-frequency words from domain texts.

The second and very important step of the proposed method is to extract relations between concepts. To solve this problem, many researchers estimate the correlation between two concepts, calculated on the basis of co-occurrence of these concepts in one sentence or one snippet [4, 5, 6]. We use a matrix that measures distance between the keywords, based on their co-occurrence in the documents. For this purpose we start with "terms-by-documents" matrix which rows correspond to keywords, columns correspond to documents and items express frequencies of these keywords in these documents. To find the pairwise distances between the keywords that are represented by row vectors in the matrix, we use the cosine measure:

$$c = \cos(\bar{x}, \bar{y}) = \frac{\bar{x} \cdot \bar{y}}{|x| \cdot |y|} \tag{2}$$

where c - is the required distance; x, y - any two rows of the "terms-by-documents" matrix, which correspond to a pair of keywords. The obtained values (distances) between keywords are measured by numbers in the range from 0 to 1. The lower distance between two keywords the smaller angle between vectors and the higher cosine measure. Accordingly, the maximum distance is 0, and the minimum distance is 1.

The quality of "terms-by-documents" matrix depends on its sparseness and noisiness. Besides the valuable knowledge about domain, the original documents contain noise which contributes to the distribution statistics. Therefore, it is advisable to decrease the noise and the sparseness

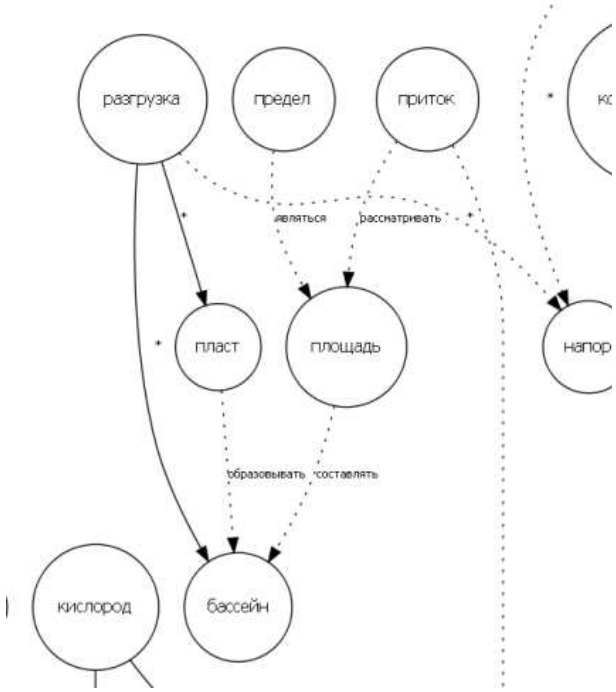

Fig. 1. - Fragment of generated concept map for the "Geology"domain

of the matrix using any factorization method. We use the method of latent semantic analysis, the essence of which is to approximate the original noisy and sparse matrix by lower rank matrix with the help of singular value decomposition [7].

The third step of the method is to transfer the extracted concepts and relations to the concept map graph. In this graph concepts are vertexes, and relations are edges (Fig. 1.).

Automatic parameterization of concept maps

As noted in [8], the use of concept maps in educational process allows better out-comes compared to traditional teaching methods. Concept maps can be used as graphic organizers of knowledge on a particular domain, and as a tool for education planning. In other words, they perform two functions: organize and represent new knowledge and monitor the current activity of the student to provide adaptive changes in the learning process [8]. The second function is very important for adaptive learning strategy, as it allows the student to construct a path of learning in accordance with his personal knowledge network. Very often in such cases, the concept maps are used as a navigation tool that guides students in the learning process. At the same time concepts can be hyperlinks to the theme of the learning course, and the relations between them can show the order of the concepts (eg, top to bottom). Obviously, using the concept maps for navigation it is recommended to reduce the number of visual elements (concepts and transitions between them).

A model of adaptive concept map proposed in this section also performs two functions: cognitive function to visualize the map of knowledge of the course, and the organizational function to adaptively manage the learning process. Also, this model is based on a number of parameters that specify students' personal models (profiles). These parameters generally take into account the individual characteristics of students, including the level of prior knowledge, the speed of their learning and perception, as well as the goals and objectives of learning. At the beginning of the learning program, students are tested to determine their knowledge on the subject. Depending

on the test results (on a scale of 0 to 10) each student obtains its own concept map. Students with a high level of knowledge get the most detailed concept map consisting of all concepts of the subject area. Students who have demonstrated a low level of knowledge obtain concept maps with limited set of concepts.

Since in our method of generation of concept maps we use Pearson's criteria to select keywords with values above a certain threshold, established experimentally, we have all the reasons to use this threshold as one of the most important parameters of a concept map. In the literature, this threshold usually is equal to 6.6, but we can vary its value to reduce or increase the list of concepts. For example, a very large value of threshold will allow keeping only the most important keywords with the highest values of Pearson's criterion, and a very small value threshold will allow selecting rare, infrequent concepts that are very specific to a narrow subject area (see Table 1).

The level of student's knowledge, test score	Pearson threshold (when the maximum number of keywords is 1000)	The number of selected concepts
1-2	25	50-100
3-5	12	100-150
6-8	8	150-300
9-10	6.6	300-500

Table 1. The scheme of selecting concepts, depending on the level of students' knowledge (for the subject area and discipline "Geology")

In the same way, we can vary the amount of relations. If among all pairwise distances to reset the values below a certain threshold, we can extract the edges connect-ing only very close concepts. In [9] it is noted that, "a reflection of relations between real objects ... involves establishing specific relationships between conceptual representations of objects in a mind". Thus, it appears that the "volume of sense perceived in the process of studying a specific training module can be estimated by the number of semantic relations between concepts in a mind"[9]. In other words, there is a direct relationship between the volume of sense perceived by student and the number of semantic relationships identified in the subject area. Thus, by varying the number of semantic relations, we can introduce another variable parameter of adaptive concept maps (see Table 2).

The level of student's knowledge, test score	Closeness threshold (cosine measure based)	The percentage of selected relations
1-2	0.9 or above	Less than 10%
3-5	0.8-0.9	10-15%
6-8	0.75-0.8	15-50%
9-10	0.5-0.75	50-100%

Table 2. The scheme of selecting relations between concepts, depending on the level of students' knowledge (for the subject area and discipline "Geology")

The initial (basic) model of concept map, constructed in this manner, will be used as a personal learning map. Each concept is associated with a link to an article or definition that reveals meaning of concept. The process of learning is built iteratively - as student learns and remembers the content of articles and semantic relations he is invited to test his knowledge. In the case of successful completion of the test a concept map is modified and updated with new concepts and relations. With this strategy, the sequence of themes under study is very important. The method of automatic construction of concept maps proposed in the previous section, does not answer the question how to number concepts for the learning, i.e. how to determine the sequence of themes under study. Now this function is the responsibility of the teacher or domain expert which determines the sequence of themes to be studied in the syllabus. But in our future work, we are going to determine this sequence by identifying the central and peripheral vertexes of the graph corresponding to the concept map.

The general architecture of adaptive learning system based on the use of concept maps

The proposed architecture of adaptive learning system is a type of "Tree of Knowledge which refers to a broad class of distributed architectures of adaptive learning [10]. This type of architecture has been created as an alternative to monolithic learning systems and supports three main modules: activities module, learning module and a module controlling student model. The learning module is responsible for the development of learning courses and learning management. However, the learning trajectory (student's activities) is stored in a distributed fashion under the control of the activities module. Therefore, the activities module regulates the flow of textbooks material to the student and is responsible for student interactions. The control module is responsible for collecting information on the implementation of student test tasks. As output data this module provides a model of student knowledge, which is analyzed by the activities module to personalize learning path.

Thus, the activities module is a key unit of this architecture, which provides adaptive learning. Usually, it supports not only the adaptive selection of learning material, but also adaptive navigation and adaptive search of new material. Adaptive navigation allows the student to learn in the most convenient way, choosing the most proper themes depending on the level of his knowledge and individual learning style. Adaptive search suggests that the system allows student to manage the search for new materials, using such criteria as passed topics, student knowledge base and the scope of his interests.

In this study, we build the architecture in which the activities module is expanded and adapted to support the concept maps. In our proposed architecture, concept maps are the basic elements that allow control over the learning process (see Figure 2).

In this study, we examined the possibility of using concept maps as the support elements of the architecture of adaptive learning systems. Such systems are designed to provide a genuine intelligence for e-learning and to organize human-computer interaction at a very high level. When designing the system, we used the "Tree of Knowledge"architecture and supplemented it with a module for automatic generation of concept maps. The novelty of our work is that we identified specific parameters of the adaptation of concept maps and suggested step by step algorithm that can be used in any e-learning system involving adaptive learning.

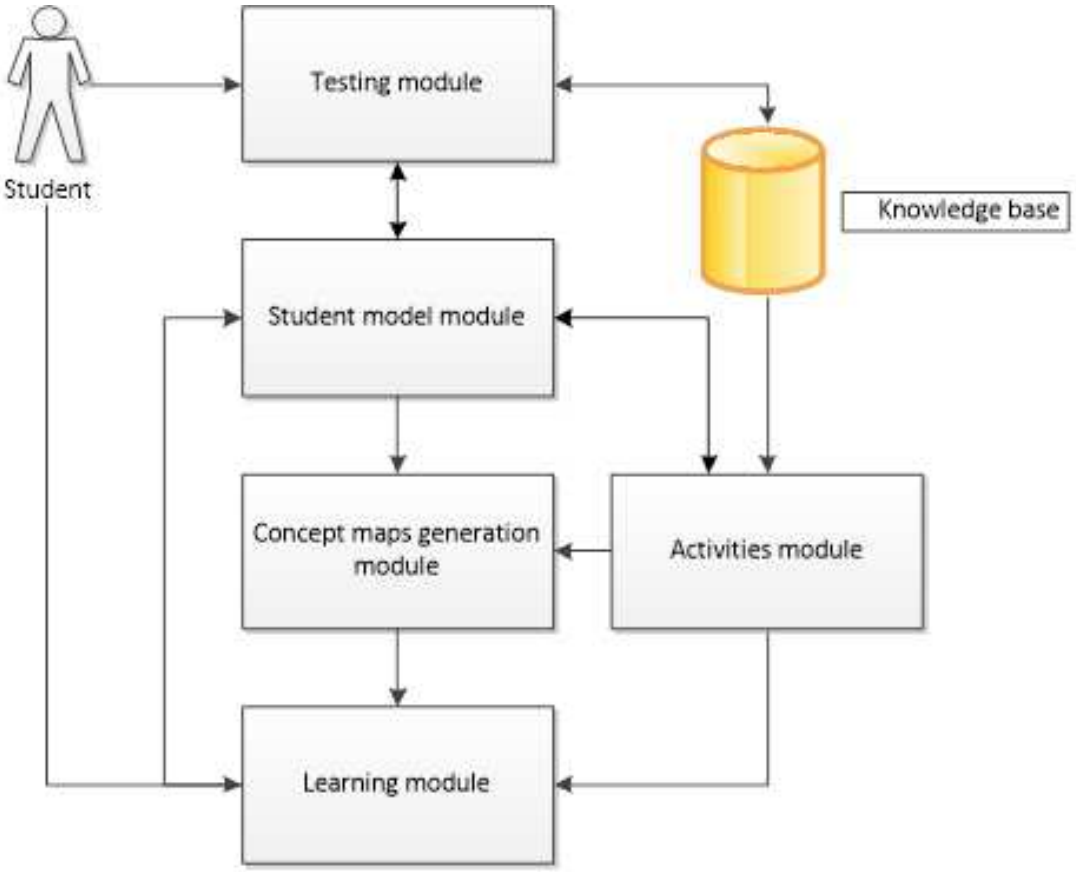

Fig. 2. - Fragment of generated concept map for the "Geology"domain

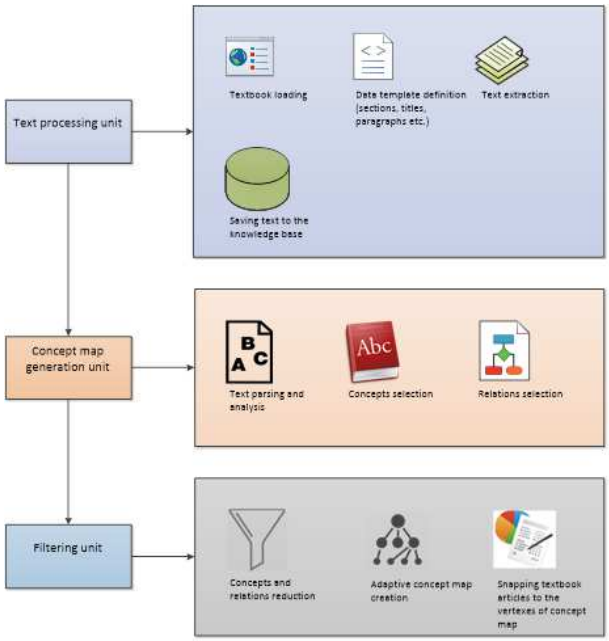

Fig. 3. - The concept maps generation module

Conclusion and future work

Figure 3 shows the architecture of the concept maps generation module, consists of three units: input processing unit, the unit responsible for construction of the general concept map of the course and the unit of construction of the adaptive concept map based on the transformation of the general concept map. This module includes the following natural language processing tools and libraries as tokenizer, lemmatizer, syntactic and morphological analyzers. In addition, it contains a unit which performs the extraction of key words (concepts) and relations.

In this study, we examined the possibility of using concept maps as the support el-ements of the architecture of adaptive learning systems. Such systems are designed to provide a genuine intelligence for e-learning and to organize human-computer interaction at a very high level. When designing the system, we used the "Tree of Knowledge"architecture and supplemented it with a module for automatic generation of concept maps. The novelty of our work is that we identified specific parameters of the adaptation of concept maps and suggested step by step algorithm that can be used in any e-learning system involving adaptive learning.

References

1. Matsuo Y., Ishizuka M., *Keyword extraction from a single document using word co-occurrence statistical information.* //International Journal on Artificial Intelligence Tools. - 2004. - T. 13(01) - C. 157-169.
2. Hulth A. et al., *Automatic keyword extraction using domain knowledge.* //Computational Linguistics and Intelligent Text Processing. - Springer Berlin Heidelberg, 2001. - C. 472-482.
3. Lott B., *Survey of Keyword Extraction Techniques.* //UNM Education, 2012. 10 c.
4. Clariana, R. B. & Koul, R., *A computer-based approach for translating text into concept map-like representations.* //Proceedings of the First International Conference on Concept Mapping', Pamplona, Spain, 2004. C. 131-134.
5. Chen, N. S., Kinshuk Wei, C. W., & Chen, H. J., *Mining e-learning domain concept map from academic articles.* //Computers & Education, 2008, 50(3). C. 1009-1021.
6. Song Q., Watanabe Y., Yokota H., *Relationship extraction methods based on co-occurrence in web pages and files.* //Proceedings of the 13th International Conference on Information Integration and Web-based Applications and Services. - ACM, 2011. - C. 82-89.
7. Deerwester, S. C., Dumais, S.T., Landauer, T.K., Furnas, G.W., Harshman, R.A.,*Indexing By Latent Semantic Analysis.*//Journal of the American Society of Information Science, 41(6), 1990. C. 391 - 407.
8. Liu S. H., Lee G. G.,*Using a concept map knowledge management system to enhance the learning of biology.* //Computers & Education. - 2013. - T. 68. - C. 105-116.
9. Alekseev V. D.,*Adaptivnoe obuchenie na osnove smyslovogo soderzhanija znanij.* //Vestnik KazNTU. - 2009. - N°4. - P. 74.
10. Brusilovsky P.,*KnowledgeTree: A distributed architecture for adaptive e-learning.* //Proceedings of the 13th international World Wide Web conference on Alternate track pa-pers & posters. - ACM, 2004. - C. 104-113.

Design and Development of Online Courses on EdX Platform

Madina Mansurova, Anna Yu. Pyrkova, and Yermek Alimzhanov

Al-Farabi Kazakh National University,
Al-Farabi av. 71, 050040 Almaty, Kazakhstan
mansurova01@mail.ru,aermek81@gmail.com

Abstract. Educational technologies, like distance learning, is developing rapidly and that development has presented educators with an opportunity to rethink and improve their pedagogical practices. Concept of massive open online courses (MOOC) become a recent and widely researched development in distance education. In the paper we introduce you to principles and practices of online course development on edX platform.

Keywords: MOOC, edX, open education, distance learning.

1 Introduction

There was an explosion of popularity massive open online courses in 2012, or abbreviated MOOC, The New York Times called this year as The Year of the MOOC [1]. It was connected with the advent of such open educational resources as edX, Coursera, Udacity and Khan Academy in the Internet. Coursera after the launch in January of 2012 to November of the same year reached more than 1.7 million subscribers – growing faster than Facebook. Many of the top level universities in the world have started to present their open courses in MOOC platforms in order to attract the best potential students and demonstrate their best courses to the wide range of users registered there. Leading universities in many countries began to develop their MOOC resources and contents, including Al-Farabi Kazakh National University.

On June 1, 2013, edX developers open sourced its entire platform. It is expected that due to the open source platform users will make it improve. Stanford University and Google became the main participants of the project with open source software, called Open edX. Google developed the site MOOC.org for everyone to post their courses, which uses the Open edX. At the end of 2013 China's universities have teamed up to organize a site for online education in China, called XuetangX. 120 high schools in France under the guidance of the Ministry of Education of France teamed up to create a national portal for online education. Queen Rania Foundation for Education and Development created Edraak – first MOOC portal in the Arab world. The International Monetary Fund has created a pilot site for online courses on economics and finance, and they all use the Open edX platform.

Here some of the features of the Open edX platform:

- molecules editor;
- code sandbox (Python, Matlab, Java);
- simulator of electrical circuits;
- Protein builder;
- interactive tables and charts;
- assessment of essays by machine learning algorithms;
- peer-to-peer assessment;
- LaTeX support;
- mobile applications support;
- analytics and other.

According to the accepted concept of development of the system of distance education in Al-Farabi KazNU, we began to implement the global experience of distance learning technologies (DLT), so we chose Open edX platform to organize online courses [2]. Currently the recent version of the system is installed and a few courses is partially downloaded. These courses are developed for students and the Virtual Academy pupils by the teaching staff and the Center of distance education. In addition to the Virtual Academy there is planned to gradually fill up content by courses for external students and students on academic mobility, as well as video recordings of open courses of summer semester and lectures of visiting professors. The main objectives of this initiative is:

- development of the DLT in accordance with modern global trends;
- gratification of the demand for high-quality educational services to the University (preparation of students, educational programs, higher education, training, retraining);
- participation of the University in uniform educational space of the international community, including the exchange and use of the full range of open resources, interaction with other members of the global educational system;
- overcoming regional barriers to learning in the University and access to education for all segments of the population;
- integration of the DLT for all kinds of forms of training in order to increase their effectiveness;
- introduction of modern tools of education quality control;
- expansion of the main activities of the University.

2 Phases of Creating an edX Course

In edX the process of creating an online course can be divided into five phases:

Phase 1: **Getting the Word Out**
This section covers what a course "About Page"is, why you need one, and what you need to provide if you want your About Page to stand out from the crowd. This section also covers additional methods of promoting your course.

Phase 2: **Creating course content**
This section covers the rationale behind the structure of the edX interface, the strategy of designing a learning sequence, the building blocks of an online course, and the necessity of creating accessible content. It also covers best practices on modes of video capture and how to structure course content to ensure learners are actively engaging with your course.

Phase 3: **Creating a grading policy**
This section covers how to establish a grading policy by defining the due date and relative weights of assignment types such as homework, labs, midterms, and final exams. It also covers the tools you and your learners use to monitor learner progress in the course.

Phase 4: **Preparing a course for delivery**
This section covers creating a course syllabus and creating a course schedule. This section also covers communication basics such as creating a welcome announcement, creating a welcome e-mail, and general guidelines for e-mails to learners.

Phase 5: **Course delivery**
This section covers course discussion moderation, verified certificates, and research and analytics.

2.1 Getting the Word Out

At the beginning to plan and design online course, it is important to develop a strategy for getting the word out about this course. An "About This Course" page (or About Page) is an advertisement for the course, composed of materials you provide to edX. The goal of the course About Page is to sell learners on why they should take this course. In figure 1 annotated course About Page is provided.

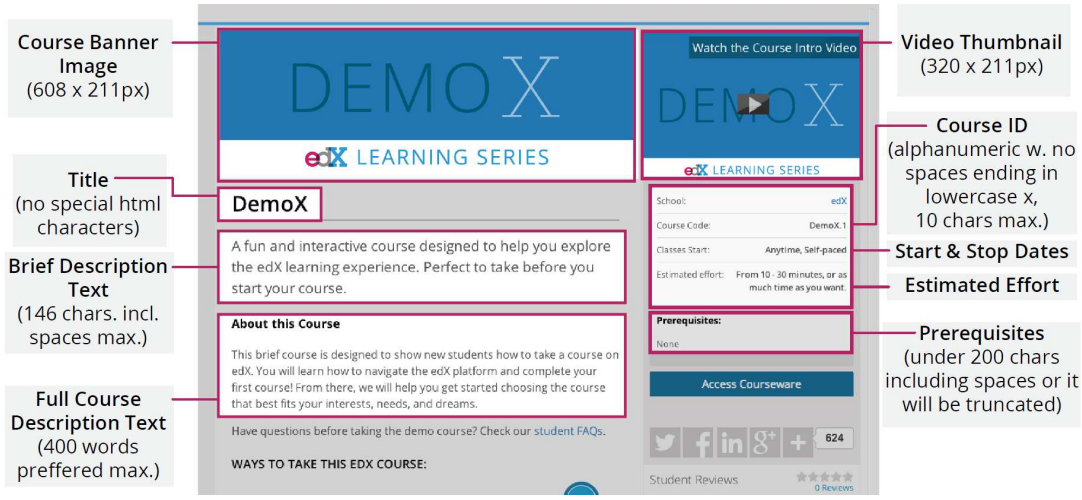

Fig. 1. Annotated course About Page

Typically, they publishes the page on the website several months before the course begins so that learners can understand what the course is about, when it is offered, and the level of effort required. This helps learners decide whether to enroll in the course.

When you prepare the text for your course description, you should ask yourself who the course is designed for and write in a tone that matches the target audience. Remember that these learners are not required to take a course as part of a major or as a required elective. They likely haven't heard from their friends which is a good course to take or which to avoid. About Page descriptions should be welcoming, energized, and—to a certain degree—provocative.

This may sound a bit vague, but think about asking questions in your About Page description to engage the reader. Avoid repetitively using "this course includes" or "assignments include." Talk about why the field or subject matter is exciting and what learners will create or be able to do, either through assignments or interactive labs.

The heart and soul of the About Page is the About video. For learners who are casually browsing edX, it is your way to stand out from the crowd and drive enrollment. Think of it as a movie trailer for your course. About videos should be 1-3 minutes long, and should make learners want to take your course. Courses with About videos have higher enrollment rates.

The About Page and About Video are two powerful ways to promote your course. There are numerous additional ways to ensure that learners are aware of and excited about your course. Use your institutional resources and networks, like institutional home page or publications such as newsletters, newspapers, and magazines; or you can use social media such as Facebook, Vkontakte and Twitter for promoting your course, driving enrollments, and engaging registered learners. Another great way to promote your course is to reach out to individuals and organizations within your professional network who may be interested in sharing news about your MOOC. For

instance, you can connect with instructors who teach related courses at your or other institutions, asking them to reach out to their students and contacts with information about your course.

2.2 Creating course content

Content is presented to learners in many ways on the edX platform, and the way learners engage with this content is the learning sequence: a set of interwoven videos, readings, exercises and discussions that cover a specific topic. The purpose of learning sequences is to promote active learning.

Interspersing exercises into the content of your course enables learners to apply the knowledge gained before proceeding to other material, thereby promoting active engagement by the learner. These exercises can be brief, ungraded questions. The value is that the instant feedback gives the learner confidence that they have understood the material they were just presented.

A learning sequence might start with a video reviewing material covered in the previous learning sequence. You might also choose to end a learning sequence with a video summarizing what you have just covered. The possibilities of how you structure your content are endless, and the choice is totally up to you. A lecture is often an hour-long exposition of a concept. Online learning exercises create an interactive experience that allows the learner to get immediate feedback on their progress through the material.

For an online course, we recommend dividing lectures into modular, bite-sized videos. Ideally, videos should be 3 minutes to 7 minutes in length and as modular or standalone as possible. Independent research on videos in MOOCs, conducted by University of Rochester professor Phillip Guo and which is summarized [3], found that shorter videos are much more engaging than longer ones. Finally, you can insert a discussion topic after each video and exercise so that learners can discuss the material with others who have also recently gone through this material. This local discussion topic, embedded in the learning sequence, will also appear in the course discussion forum, which we will cover later.

There are four main content building blocks you can use to build your course. These may not sound like a lot, but within these four building blocks there are a wide range of possibilities (fig. 2).

Fig. 2. Component adding menu

HTML: This is both the most basic content type and the most versatile. You can add rich text, images, links, and more.

Discussions: Allow for learners to engage in course discussions that are embedded within the natural flow of a course.

Problems: edX offers a wide variety of out of the box problem types, as well as the ability to create and customize more advanced problem types.
Videos: edX's built-in video player possesses many features to optimize the learner experience.

2.3 Creating a grading policy

Establishing the grading policy for your class is an important aspect of course design. The edX platform gives you considerable control over the exact nature of the grading policy. As you can see below, you can adjust grade ranges, change the names of grades (for instance, you can change a letter grade of "A"to the word "Excellent"), and decide on the number of possible grades.

It is important to note though that if your course offers students the opportunity to earn a certificate, any student earning a grade over the "F or "Fail threshold will qualify for a certificate. This is true regardless of how many grade levels you add in the grade range. Thus, when establishing your grading policy it's important to consider where you want to set the bar for receiving a certificate.

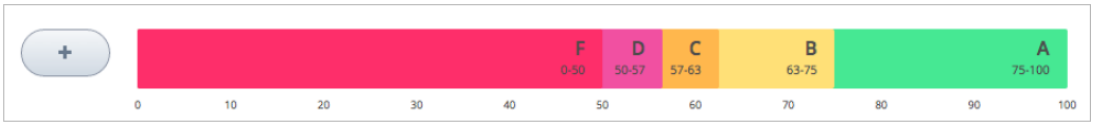

Fig. 3. Grading scale

For each graded assignment in your course, you can specify a release date and time when learners can first access the assignment, as well as a due date and time at which point the assignment is closed. You also have the ability to apply a grace period on all assignment due dates in order to give students some additional leeway.

2.4 Preparing a course for delivery

Much like in an on-campus course, the preliminary course material is the logistical scaffolding of an edX course. These are a set of introductory materials that learners view when they first enter a course, and are not unlike the administrative material you might hand out on the first day of a campus class.

Preliminary course material can include:

- A syllabus
- A course schedule
- Any specific collaboration or course discussion guidelines
- A welcome announcement on the Course Info page and future announcements
- A welcome email and future email communications

The Course Info page is the first page that learners see when they access your course. It's a very good idea to create a welcome announcement on the Course Info page to serve as a landing pad for new learners and to provide guidance and instructions about the first steps learners should take. The welcome announcement should encourage learners to check out the syllabus page, inform them about course discussions, and tell them what is available for viewing. Some instructors have also pointed to a self-diagnostic quiz and material to help learners make sure

they're prepared for the course. In addition to a welcome announcement at the beginning of your course, we encourage you to post a regular stream of announcements on the Course Info page throughout the course.

While your Course Info page serves as home base for disseminating information directly to your learners, we recommend that you also send crafted e-mails before and during the course. To ensure that your learners are engaged and responsive, we recommend that your e-mail communications be brief and to the point. Personal e-mails such as these will increase the likelihood learners will complete your course. Usage may vary, but we recommend outreach to learners two months, one month, one week, and one day before your course begins.

2.5 Course delivery

The final phase of creating and launching an edX course is course delivery. Course delivery includes actively monitoring discussion forums, communicating with learners via e-mail and other channels, monitoring student activity and performance, and awarding students certificates.

Student collaboration through course discussions and other channels are an integral part of the MOOC experience. Yet not all learners have experience participating in online discussions and other forms of online collaboration. Offering guidelines for collaboration encourages your learners to help each other learn. Likewise, discussion guidelines can give your worldwide audience an idea of the types of conversations you would like to foster in the forums.

The edX Insights feature makes information about courses available to course team members who have the Course Staff or Instructor role [4]. edX Insights provides these course team members with data about learner backgrounds and activities throughout the course. Using edX Insights can help you validate the choices you made in designing your course. It can also help you re-evaluate choices and inform efforts to improve your course and the experience of your learners.

3 Conclusion

Creating online course is a very complex and interesting process that requires careful structuring of the course material and design of high quality content. In the near future, online courses can be an important part of the educational process in secondary and higher educational institutions. In these circumstances, all teachers will have to learn the skills to create online courses that will help them to find students worldwide.

References

1. The New York Times, http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html
2. MOOC in al-Farabi KazNU, http://edx.kaznu.kz
3. How MOOC Video Production Affects Student Engagement, http://blog.edx.org/how-mooc-video-production-affects/
4. edX Insights home page, https://insights.edx.org/courses/

Positive Practice in the Implementation of Moodle in E-Learning

Petrovic V. Vera[1] and Jokic Anja[2]

1 - School of Electrical Engineering and Computer Science of Belgrade,
2 - Faculty of Natural Sciences, Pristina University with temporary place of residence in Kosovska Mitrovica

Abstract. In line with the development of information and communication technology, today education is not limited to only what required is. Every contemporary individual wants to improve and learn as much as possible and beyond. That is exactly how the popular term of lifelong learning originated. In order to survive in a variety of information and knowledge, it is necessary that one constantly improves himself and keeps pace with the world. As one of the ways to meet these needs, the e-learning was developed, which is progressively gaining momentum and popularity in the world. E-learning brings for sure certain number of advantages in the educational process. It is not an alternative to the existing educational process, but rather an integral part of it, its expansion and improvement. With the introduction of e-learning, have the role and the importance of teachers as mentors, coordinators and participants in the educational process grown. E-learning enables students to be in the center of the educational process, and to take an active role and responsibility for educational outcomes. E-learning is certainly a high quality education process, in which all actively cooperate with the aim of achieving the learning goals that have been set. Moreover the modern information and communication technologies are intensively used to create an adjustable virtual environment. Moodle is offered as one of the good software tools. This paper discusses the application software package Moodle in the process of e-learning on subjects that include content courses basis of informatics and computer science, and studied at the undergraduate level of higher education.

Keywords: E-learning, Moodle, Electrical.

1 Introduction

Development of information technologies and constant innovating of educational technologies results in alterations in methods and forms of educational activities, including the organization which would appear optimal in the era of mass implementation of internet and electronic sources of knowledge. Education, as a rule, gradually embraces new technologies in view of production, traffic, service providing. However, multimedia systems, distance education, virtual schools and other technologies which result in increase of students engagement, higher quality of knowledge assessment and students advancement in accordance with individual abilities and prior knowledge, were introduced in education [1,2]. In addition to favoring knowledge as principal global resource for the future, significant efforts were invested in developing and expanding of all forms of electronic support to educational process, during the last decade. Alongside utilization of new technologies in enhancement of formal education, distance education was also developed, with basic goals to set more flexible infrastructure, and with it the availability of this form of learning to each student, to alleviate the overall level of digital literacy in academic population and to develop top quality educational content. Education supported by information technologies implies three primary components at least:

1) Computer Assisted Learning CAL
2) Computer Assisted Research
3) Distance Learning - DL

Computer Assisted Learning is used most frequently and it is highly convenient for establishing interactions between students and computers in order to improve existing learning technologies

and make lectures more obvious, dynamic and interesting. Computer Assisted Learning includes multimedia educational software, computer simulations, virtual reality, artificial intelligence. Utilization of information technologies envisages individual gaining of knowledge, constant return of information and monitoring of students progress which helps teacher assess students knowledge more realistically and points them to other didactic media so that they could gain new knowledge without difficulty. Computer Assisted Learning is, today, used extensively at high education institutions for theoretical researches of literature from different areas and for empirical researches with using of adequate statistical software. Theoretical research of literature is almost inconceivable without the use of computer technology, because, today, almost all major books, papers, studies and repertories from expert and scientific conventions are transferred to electronic form and placed on web portals of publishing houses, faculties and libraries. Distance learning by using of computers, telecommunications, cable television is more frequently implemented in education. Distance education represents an instructional method of work with students which does not require presence of students and teacher in the same room. Development of distance education dates from mid 20^{th} century when radio broadcasts and pen-pal schools were used in order to provide education to students with no means of travel to school or simply unable to attend lectures. On a basic level, distance education is established when both teacher and student are physically separated by long distance and when technology (i.e. speech, video, data and printing) is used to abridge this gap. These types of educational programs can offer second chance for attaining faculty level education to adults, can reach those who are troubled with time constrains, away from faculty centers or with physical disabilities, and can also refresh their knowledge and improve skills in accordance with development of science and technology in their related area of expertise. This paper contains an overview of Moodle software package in distance education at School of Electrical and Computer Engineering of Applied Studies from Belgrade in course subject Basics of information technology and computer science and a proposal for implementation of such method of work in course Theoretical basics of information technology on Faculty of Natural Sciences University of Pristina with temporary place of residence in Kosovska Mitrovica.

2 Moodle software package

Moodle software package is used for creation of internet courses and websites and it was created as support to electronic learning. Term Moodle is an abbreviation of Modular Object Oriented Dynamic Learning Environment. It is created as open source software, with open source code. It essentially means that author of this software package has protected its rights but whoever uses it has additional rights. Basic characteristics of this software package, as support to distance education, can be classified into several categories: High availability ability to serve thousands of users simultaneously;
Scalability ability to absorb increase of users without a drop in performance;
User friendly ability to have the user (student or teacher) quickly learn how to use the system;
Interoperability ability to integrate in the existing software in related institution;
Stability stabile version of Moodle software provides uninterrupted services to student and teacher population; Security - characteristic of system to pose security risk not greater than any other component of information system in related institution.

3 Course Electrical materials and components

In course subject Electrical materials and components, studied on the first year of School of Electrical and Computer Engineering of Applied Studies from Belgrade, an online course was

organized as additional support to traditional methods of providing education. Thanks to Moodle software package, during this course, students have access to learning materials, complete set of lectures, announcements sent on email via forum, contact and discussion with professor and assistants and can test their knowledge by taking related tests. After entering the URL `http://www.wikibooks.org` in web browser, the home page of Moodle opens, as shown in Figure 1. This home page contains all distance education simplified courses in Moodle environment at Electrical Engineering and Computer Technology College of Applied Studies from Belgrade.

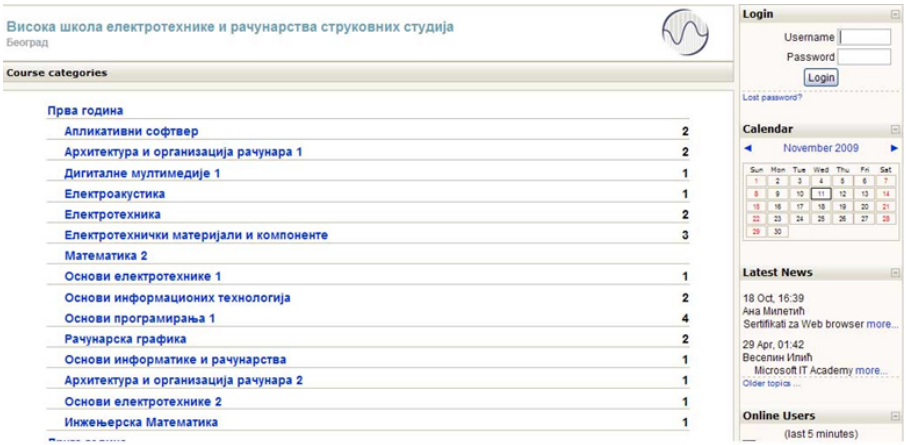

Fig. 1. Home page of Moodle system

Titles of courses at School of Electrical and Computer Engineering of Applied Studies (SECEAS) available to students through distance education represent links (hyperlinks), as shown in Figure 2.

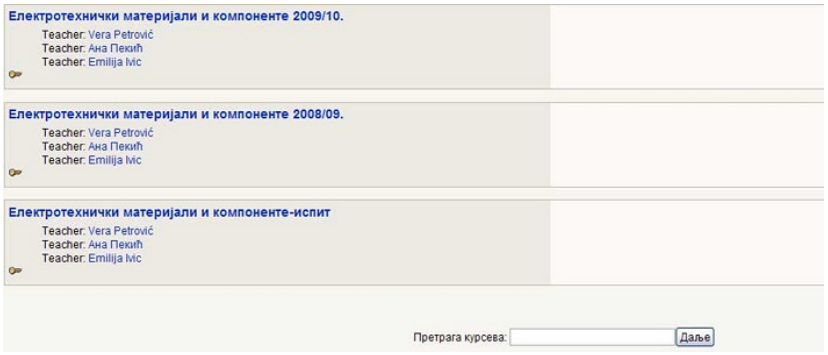

Fig. 2. Course subject page

By clicking on a related link (course subject title) and entering appropriate access password, Figure 3, user logs in and becomes an attendee of the selected course, in this case Electrical materials and components course.

After students pass the procedure of updating their profiles, and it is necessary for a student to receive a password by school network administrator, which will then be used during their entire schooling in SECEAS, Moodle system sends an activation link to the email addresses they

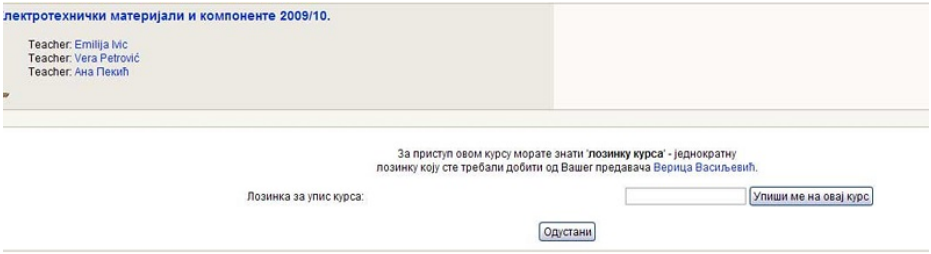

Fig. 3. Enrolling (logging in) to course

had entered. After the link for account activation is activated, Moodle page for accessing the system opens. By selecting the course Electrical materials and components on attendees home page, Figure 4, students enter their identification password.

Fig. 4. Enrolled course attendee home page

Should someone like to see the content of one topic on a page in any moment, a click on little square on the right side is required □. After clicking, the view of the same topic remains almost identical but all other topics are invisible. In order to make the topics visible again, a click on the double square in the upper right corner of the topic is required. In case someone would like to see all topics again, clicking on the double square which will appear on that spot is required. This enables simpler approach to specific topic in cases when the course has many topics, Figure 5.

Moodle system also enables testing of course attendee knowledge by means of online testing. At Basics of information technology and computer science course, students check their knowledge related to laboratory exercises prior to each exercise. Taking the test is conducted by clicking on the test title within certain topic, after which a web page opens, as shown on Figure 6. Information concerning the test will appear on the first page during opening of the test.

– Number of attempts
– Time limitation
– Grading method (first attempt, last attempt, average grade and highest grade)
– Duration of the test

Fig. 5. Topic overview on the page

– General characteristics of the test and related information (purpose)

Fig. 6. First page of the test

Student starts the test by clicking on Start test button and before he receives questions Moodle systems informs him on the allowed number of attempts for solving the test, time limitations (if any) and any other information. Starting from the moment the testing is commenced, a clock counting down the time to test closure is activated. When time expires the test shuts down and submits itself. When student marks the answers he believes are correct, in order to complete the test and receive return information from the system, clicking on the button Submit all answers and complete the test as shown on Figure 7 is required. Besides this option there is also an option to save without submitting, which memorizes all answers from the test but does not stop the clock.

When a student completes the test, a web page with return information regarding the scores on the test will appear, as well as a possibility to see where mistakes were made, information on the remaining number of attempts and scores from previous attempts.

Fig. 7. Button for submitting test answers

4 Conclusion

Distance learning is projected to have the content follow the educational program of related subject courses, but using hyperlinks also provided opportunity for in-depth researches in particular areas, which suits individual interests of students and all others who would use these didactic materials for formal or informal education. It is necessary to establish permanent evaluation of distance education and its enhancement in accordance with changes that occur in the developed parts of the world, as well as on the basis of experiences and opinions of lecturers and students in practice. Public approach to web content stresses the need for protection of privacy of students information, so a special database related to students with password protection and availability to authorized personnel only would be of necessity. Database containing results of students progress must be protected and archived on CD-s, not only for protection of privacy but also for preventing eventual misuse by the side of students. Nevertheless, in this early phase, assessing the students knowledge by means of distance education would only serve as additional information to lecturers, not as the only one, and surely, during its use, another organization adjusted to social environment we live in will be imposed. However positive experiences from this course may successfully be implemented onto other high education facilities.

References

1. Milosevic H.,Petrovic V., *Informacione tehnologije kroz softversku platformu Moodle*,MIT 2011, Vrnjaka Banja Budva, (2011)
2. V. V. Petrovic, A. Pekic, E. Ivic, *CDaljinsko obrazovanje primenom softvera Moodle* ,YUINFO 2010, Kopaonik, (2010)
3. Vera V. Petrovic, *Primena daljimskog obrazovanja u Visokoj skoli elektrotehnike i raunarstva*, INFOTEH 2010, Jahorina, (2010).

Structuralization of Categories of the "Knowledge"Pedagogical Science in the Process of Informatization of the Society

B.K. Shayakmetova[1], G.T. Omarov[2], Sh.Ye. Omarova[2], N.T. Orumbayeva[1]

[1]E.A.Buketov Karaganda State University,
[2]Kazpotrebsoyuz Karaganda Economic University
Karaganda, Kazakhstan
{kazahzavod,omarov-gali,sheo_1953,orumbayevan}@mail.ru

Abstract. In the present work the concept of "knowledge"is examined in the pedagogy in higher schools. Today we have to spend time fighting for the quality of students ' knowledge. To transfer knowledge to students means to form the essence of concepts, to help to understand the natural connection of phenomena of material reality and to clothe it in a correct and clear verbal or other form.First of all it is necessary to ensure that the student has clear, specific representation of the object before quoting a particular phrase.

Keywords: knowledge, pedagogical process, category, quality of knowledge, study material, career guidance, formalism, stencil, information disciplines, theory.

It is difficult, maybe even impossible, to give a clear definition of "knowledge": firstly, this concept is one of the most common, and as such it is always difficult to give an unambiguous definition; secondly, there are many different types of knowledge, and they cannot be put in one row. Knowledge is the result of cognition of reality tested in practice and certified by the logic, reflected in human consciousness in the form of ideas, concepts, opinions and theories. The following keywords are really important here: Certified by logic - i.e. logically consistent; Tested in practice - i.e. whatever the theory, it can be recognized only when the theoretical calculations agree with experimental results. Hence, there is such property, as the objectivity of the results of the experiment which must be independent of the identity of the experimenter (emotions, prejudices, personal likes and dislikes, etc.). Knowledge is the form of existence and systematization of the results of cognitive activity. Knowledge helps people to organize their activities and to solve various problems arising during the process. Knowledge in a broad sense is a subjective image of reality in the form of concepts and notions. Knowledge in the narrow sense - the possession of verified information (answers to questions), allowing to solve the task. Knowledge (of the subject) is the confident understanding of a subject, ability to deal with it, to understand it and to use it in order to achieve the intended objectives. Knowledge in the theory of artificial intelligence and expert systems - a set of information and rules of inference (the individual, society or the AI system) about the world, the properties of the objects, regularities of processes and phenomena, as well as the rules of their use in order to make decisions. The main difference between knowledge and data consists in their structure and activity, the emergence of the basis of new facts or new relationships can be a source of changes in decision making. Knowledge is captured in the images and marks of natural and artificial languages. Cognition is not limited to the field of science, knowledge in one or another form exists outside of science. Each form of social consciousness: science, philosophy, mythology, politics, religion, etc. - correspond to specific forms of knowledge. There are also forms of knowledge with conceptual, symbolic or art-model basis [1]. There are different kinds of knowledge: scientific, non-scientific, and mundane-practical (common sense), intuitive, religious, etc. Scientific knowledge is the type knowledge based on rationality, characterized by objectivity, universality, and claiming

the common significance. Scientific knowledge is the process of obtaining an objective, true knowledge. Its task is to describe, explain and predict the process and phenomenon of reality. The scientific revolution occurring in the course of development of scientific knowledge leads to the change of theories and principles, followed by periods of normal development of the science (deep and detailed knowledge)[1]. Scientific knowledge is inherent to logical validity, verifiability, reproducibility, the desire to eliminate mistakes and to overcome contradictions. According to the degree scientific knowledge can be scientific and non-scientific.

Scientific knowledge can be divided into:

- empirical (based on experience or observations)
- theoretical (based on the analysis of abstract models).

In any case scientific knowledge must be grounded on empirical or theoretical evidence base. Theoretical knowledge - abstractions, analogies and schemes showing the structure and nature of the processes occurring in the subject area. This knowledge is explaining the phenomena and can be used to predict the behavior of objects. Theoretical ideas arise on the basis of generalization of experimental data. At the same time they affect the enrichment and the change of empirical knowledge. Theoretical level of scientific knowledge involves the establishment of laws, giving the opportunity of idealized perception, description and explanation of empirical situations, i.e. to understand the nature of phenomena. Theoretical laws are more strict, formal, if to compare them to the laws of empirical knowledge. Terms describing the theoretical knowledge are idealized and abstract objects. Such objects cannot be subjected to the direct experimental verification [1]. Urgency of the research shown below is identified by the need to enhance the pedagogical process in the context of continuous education and structuralization of the "knowledge"pedagogical category in the process of informatization of society. Structuralization of the main category of the "knowledge"pedagogical science is based on two fundamental assumptions: firstly, the entire system of knowledge can be hierarchically situated; secondly, such a streamlining should be conducted in accordance with the evolution of knowledge systems, i.e. by the production of natural classification schemes. The subjects of research are the information needs of the individual and the structure of education system [2]. The research raised the hypothesis about the possibility to synthesize the informative education system, if cognitive activity is based on hierarchically organized knowledge, acting as a system, providing and stimulating the needs of the individual. Social development of society is characterized as a complex dialectical process, the effectiveness of which depends on the level of development of professional adaptation. One of the ways to address these social-pedagogical problems is the professional orientation of students, which in turn determines the social adaptation of young professionals being a part of the problem of upbringing. Now we have to spend time fighting for the quality of students ' knowledge. So, in order to convey the knowledge to the students we need to form understanding of things and phenomena in their minds, to reveal the essence of phenomena - to form concepts, to help them understand the natural connection of phenomena of material reality and to clothe it in a correct and clear verbal or other form. One of the drawbacks that affect our education is the formalism in knowledge of students, which should be comprehended as:

a) separation of form and contentof expression;

b) rote memorization of the study material without clear understanding.

In computer science formalism in knowledge of the students is expressed in:

1. The absence of consciousness, in the uptake of computer science and in the predominance of the external form over content. For example, students freely characterize the Windows operating system but they are hardly describing the Unix.

2. The predominance of memory over understanding. For example, at the little change in the task and designation of the initial data in the problem book students are often get lost and can't create a program for this case.

3. The predominance of the stencil, template. Students are not always looking for the most beautiful, convenient way of solving the problem and try to reduce their decision to a familiar model.

For example, at the creation of a local area network (LAN) they use the ATM (Asynchronous Transfer Mode) network technology, considering it more modern and frequently used, while X. 25 networking technology would be preferable.

4. The separation of theory from practice. For example, students quite easily create programs (with appropriate training) for tasks and repositories; however, if the task is taken from the practice they start to face problems.

In order to avoid the formal knowledge of students it is necessary to:

1. Pay special attention to the precision and clarity of forms of expression, to the explanation of the meaning of each word expressing the fact of information.

For example, at the characteristic of the ATM (Asynchronous Transfer Mode) network and network technology occurs the phrase: VPI and VCI identifiers are used to indicate ATM virtual connections. Students often don't realize why the words "IDs "VPI "VCI "virtual connection they never think that the ATM network and network technology can't be defined without the words "IDs "VPI, VCI"and "virtual connection".

First of all it is necessary to ensure that the student has clear, specific representation of the object before quoting a particular phrase.

The correct notion of special terms is impossible without understanding the relationships between them, without understanding their mutual dependence. The student must be aware clearly that each definition or rule is valid only for that class of object for which it is installed.

However, it is also important to take care of work on the form of expression, knowledge shouldn't be separated from working on content of knowledge; work on the expression of knowledge must be in organic connection with the conscious assimilation of the content of educational material.

2. To make students to master the material consciously. It is often observed that the student reproduces the covered material mechanically, without sufficient understanding. The student does not figure out the core material and the conclusion because of misunderstanding of presented material. We also know that sometimes students when viewing the material skip the intermediate portion of reasoning or they start to explain one material, finally ending with another.

In order to overcome these drawbacks, it is useful to:

a) change the presentation of material, for example, to expound it, changing the order, changing the angle of view, unless if thoughts for this case are true.

By changing the graphic support, we encourage students to understand the essence of the material.

b) when presenting the material, we should try to ensure that the student understands why there are one or more graphical reasoning. It is necessary to oblige the student to explain the usefulness of this graphical reasoning.

It is also necessary to ensure that the student could apply the general rule for each particular case.

3. To make students understand the objectives of studying of each topic. It is useful to begin its presentation with a question. We should not think that the question is only the listing of content. The students should be set a specific goal, indicating the value of the topic for further

study of computer science and information disciplines, there is also should be shown the practical value of the new material, the connection with the previous material, etc.

For example, starting the topic of network technologies it is necessary to indicate the importance of this issue, to enumerate all the possible cases of application of these technologies, to enumerate the advantages and disadvantages, and only after that steps to proceed to the presentation of material on local computing networks. The students in this case will have a certain perspective and clarified scope of the topic.

Sometimes it is useful to issue the brief historical sketch before the start of topic. However, it is necessary to save the time, because in some cases such a preparatory work can take too many hours, if it will be sufficiently deployed.

It is useful to review the learned material and to show its practical application after each topic.

4. To teach students how to apply their knowledge at the solution of practical problems. Therefore, it is useful to offer students the chance to create a local network, for example, with the neighbors on the floor or around the house.

5. When solving problems the template should be avoided whenever it is possible. It is often observed that the student solves the problem, with the tacit consent of the instructor, i.e., proceeds to the solution of problems, so to say "blindly not realizing why he is doing it in this way. Moreover, the student is unable to explain and justify his actions. Often the student is trying to reduce the problem to some previously established type, not looking for the most simple or original solution. All of this contributes to the formal assimilation of computer science and information disciplines. Therefore, it is necessary make the student to form the layout of solving problems, to justify each successive stage and to point out what he wants to achieve.

6. While asking the student it is not right to be satisfied with the response of the student, without figuring out whether the student understood what he is saying.

Therefore, after the presentation of the material to the student, it is useful to offer him questions, the answers to which will show the real understanding of the material.

For example:

1. Concept of algorithm. Properties of algorithm.
2. Describe the forms of algorithm presentation.
3. What is the operator scheme?
4. Give examples of the types of computingprocesses.
5. What kind of algorithmic processes are most commonly usedat the programming?
6. What is the content of problem statement?
7. To characterize the stage of tasks implementation.
8. What is the solution of the problem? Such questions promote the deep understanding and conscious relation to information provisions and therefore will contribute to eradicating of the formal knowledge of the students [3]. The Professor Sh. Ye.Omarova [4], indicates that the formalism in computer science and information disciplines sometimes mixed with the requirements of formal logical speed of its presentation. The fight against formalism is understood as the struggle for the expulsion of the requirements of formal logical rigor justification of various truths from the process of education. Such an understanding of the struggle against formalism is fundamentally wrong. The expulsion of rigor of reasoning, necessity to justify each conclusion, from computer science and information disciplines, will lead to the most disastrous results. If a University student will not understand the need for strict reasoning, will not appreciate their advantages, will not be able to apply this reasoning to the resolution of some problems, then the student will really develop a formal approach to the subject. If the preparatory work would not be carried out (indication of the feasibility of introducing some of the concepts, justification

of every concept from the practical point of view, demonstration of their importance, etc.) in an abstract and rigorous presentation of computer science and information disciplines, students create a belief that the subject of study is the arbitrary number of non-justifiable conditions; however, the most inquisitive students will protest against these conventions, and perhaps even will offer their own rules. As mentioned above, the formalism in knowledge of students is a major weakness. Students with formal knowledge are not prepared to practice. To overcome the formalism in knowledge means to prevent the separation of words from thoughts and to make the idea meaningful. Only when the mind of the student correctly reflects the reality and such a reflection will be expressed in correct and clear form, the formalism in students' knowledge will be completely eliminated. We also note the following provisions affecting the learning of new material and memorability of the material which is already passed (how long the material remains in the memory), provision on culture of informative speech, because the competent speech solves the issue of informative development and knowledge of students. Proper informative speech is expressed in the correct writing of terms, in the knowledge where these terms can be used, and in understanding of the meaning of terms and special informative expressions. To teach the correct writing of information terms is the task for teachers themselves. Spelling of the newly introduced term should be displayed on the board and students should write them in their notebooks. The right stress of words is also important. The correct abbreviation of words in the information records is important too. The lack of uniform requirements for the abbreviation of words leads to negligence in reducing and promotes the illiteracy. There is some special reference of standards of abbreviations; it is necessary to use it. Carelessness and inaccuracy in the use of terms leads to misunderstanding of information laws, to confusion in the formulation and practical application of them. It is needed to check, do students associate what they say with real facts, do not they just repeat the formally memorized sentences. Requirements of the teacher to students to present their minds fully and coherently at any kind of theoretical explanation of the procedure and solution of problems contribute to the development of speech information. The teacher should require from student not only the correct wording, but also cleansing of speech from words-parasites, or the terms they invented. It is necessary to teach students to listen to their speech and to follow the alignment of words in sentences, because the improper coordination of information in speech is distorting the thinking that often leads to gross errors in the conclusions. The source of these errors, inaccuracies, nonsense from a logical point of view in the speech of students is primarily occur from the methods of teaching the computer science in the secondary school. Teachers often justify their lack of attention to the speech that the study of correct speech is a matter of Russian language teachers, and in the best occasion they correct mistakes mechanically. We need to understand the following situation: the process of expression of ideas by students is in the nature of dialectical interaction; speech is forming the ready conscious thoughts, but at the same time thoughts are shaped by the speech. Therefore, each teacher, including the information disciplines teacher, must work routinely and systematically towards the shaping of the thoughts of students on the elaboration of correct forms of expression through speech and writing. The following ways for the right and competent speech:

- First of all teacher's speech must be correct and serve as the model to students. Therefore, the teacher should use right formulations, to plan the records on the blackboard, which will be written in notebooks.
- Teachers must overcome the habit of putting the students at the polling question after question and be content with fragmentary answers. It is necessary to provide the student with the opportunity to express their thoughts fully, only with occasional remarks: "why, think is it true to put them in the right direction using these replications. In addition to

asking questions of the material, perhaps, to offer to the students to orally repeat some arguments, definitions, concepts; it cannot be assumed that the student can only express the theoretical reasoning through writing. Verbal reasoning, besides the fact that contributes to the development of spatial imagination, also develop and consolidate the oral and written language of the student.

- It is extremely useful to specially stop and to analyze sentences and formulations where the particular word is missed, to research what leads to the omission of the words, and thus to educate the students to understand the significance of each word in these sentences.
- During some control works it is necessary to include theoretical questions that are needed to be answered or explained in the text.

It is necessary to explain them from the beginning what kind of relationship exists between the general theoretical position and private judgments. The student must clearly imagine that the adoption of any provision on the basis of several particular cases, even a very large number of them, has no credibility. On the other hand, the statement that the experience, understood in the broadest sense of the word, should not be given the importance is incorrect. Experience is a must and it is the starting point for many theoretical considerations, although the experience may not be used unconditionally for the dissemination of some information on all possible cases. It is also necessary to show students that the theory helps us to study the phenomenon, to avoid unnecessary experiments or consideration of particular cases, which consume time and do not cover all the cases. Training of the carefulness, attention and responsibility is also important while teaching the information sciences. It is necessary to decisively deal with carelessness, irresponsibility while performing work. The teacher should require from students the good and cultural performance, both on the board and in their notebooks.

References

1. Philosophy for graduate students: study guide / V.P. Kokhanovskiy [andothers]. – 2ndpubl. – Rostov : Fenix, 2003. – p. 448 – (Higher education). – ISBN 5-222-03544-1
2. Shayakhmetova B.K., Antipov Yu.N. On particular questions concerning the solution of questionsoccurring at thesolution of problemsof informational character.Materials of the international scientific-practical conf. November 20, 2010 – Moscow-Kaliningrad-Smolensk. – T.2. – p.101-106
3. Shayakhmetova B.K., Omarov T.E. On the proposed approaches to improvement of the content of education of the information systems specialist // Karaganda university herald. "Pedagogics"series, 2005 No.4 (40) – p.92-95
4. Omarova Sh.Ye. Information science: Study guide. – Karaganda, 2006. – p.196

Database Design for the Sectoral Frame of IT Qualifications Within TEMPUS Project ”QUADRIGA”

Baidaulet A. Urmashev, Anna Yu. Pyrkova, Madina E. Mansurova, Erlan P. Makashev, Aimurat Zh. Burlibayev and Magzhan S. Sarsembayev

Al-Farabi Kazakh National University, Almaty, Kazakhstan
{Baidaulet.Urmashev,Anna.Pyrkova}@kaznu.kz,{mansurova01,Makashev_Yerlan,Aimurat06,magatrone}@mail.ru

Abstract. One of priority objective of Tempus project "QUADRIGA"is development of the project of a sectoral frame of qualifications (SFQ) in the field of information technologies on the basis of the All-European recommended frame and taking into account experience of the European and Russian higher education institutions.
Implementation of the project will allow to provide a basis for modernization of educational programs, to promote closer cooperation and mutual understanding between universities and employers (the enterprises, industries, public sector) that will be useful to graduates of higher educational institutions of Central Asia and will allow to strengthen the Centers of Career in higher education institutions of Central Asia for expansion of opportunities of graduates in employment and creation of high-quality employment.
Today participants of the project developed the project of a database for a sectoral frame of qualifications taking into account influence of professional competences on formation of the contents, the purposes and problems of disciplines of higher educational institutions.
Thanks to the developed structure of a database it is possible to track interrelation between the labor functions, necessary for successful work in sector of IT technologies, and educational competences of discipline; to see and estimate process of formation of professional competences of the graduate of a higher educational institution; to define the education level demanded for assimilation of necessary skills (undergraduate studies, magistracy, PhD doctoral studies), etc. In the future the team of the QUADRIGA developers will make an attempt to use this database for development of an educational program of IT specialties.

Keywords: Sectoral frame of qualifications (SFQ), database, educational program of IT, employers, competences.

Introduction

Since 2012 al-Farabi KazNU, KIMEP, EKSTU named D. Serikbayev, and the Karaganda state industrial university work on international Tempus project "QUADRIGA". One of priority objective of this project is development of the project of a branch (sectoral) frame of qualifications in the field of information technologies on the basis of the All-European recommended frame and taking into account experience of the European and Russian higher education institutions.

The main direction of reforming of the higher education of the Republic of Kazakhstan is creation of conditions on ensuring its adaptation to changes in economy, and also to entry of an education system into world educational space. It is known that the Republic of Kazakhstan is the first Central Asian state which in 2010 joined the Bologna declaration and became the full participant of the European educational space. The purpose of participation of Kazakhstan in Bologna Process is expansion of access to the European education, further improvement of quality of education, and also increase of mobility of students and teaching structure [1].

Now in the Republic of Kazakhstan the State program of a development of education for 2011-2020, which purpose is cardinal modernization of an education system, significant and steady increase in investments into education, improvement of its quality and an exit to the European level, is realized.

National and sectoral frameworks of qualifications

One of the actions directed on implementation of this program is introduction to the Labour Code of the Republic of Kazakhstan the addition in the form of chapter 10-1 "National system of qualifications"("About modification and additions in the Labour code of the Republic of Kazakhstan"from February 17, 2012 No. 566-IV LRK). According to [2] National System of Qualifications (NSQ) is a set of mechanisms of legal and institutional regulation of demand and offers on qualification of experts from labor market (fig. 1). The purpose of introduction of NSQ is creation of flexible system of qualifications in RK on the basis of effective mechanisms of legal and institutional regulation of interaction of professional education and labor market.

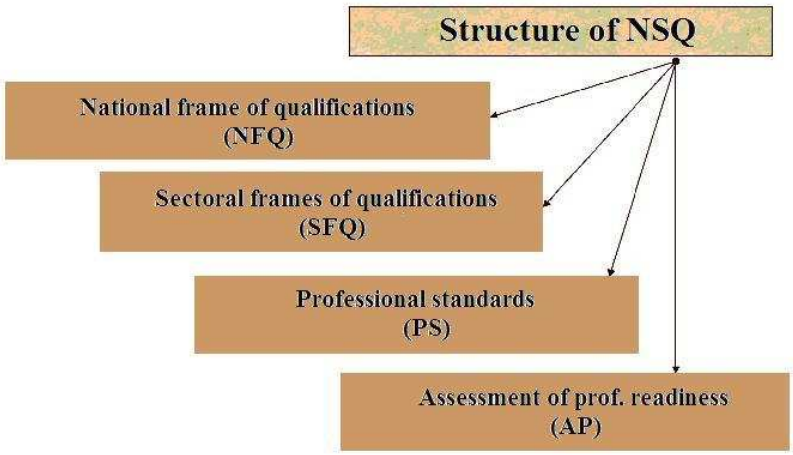

Fig. 1. Structure of NSQ.

The main terms and concepts used in the National Frame of Qualifications (NFQ) and presented in the document "A national frame of qualifications. General provisions"following:

1) A national frame of qualifications – the structured description of the qualification levels recognized in labor market;

2) The Branch Framework of Qualifications (BFQ) – the structured description of the qualification levels recognized in branch;

3) The Professional Standard (PS) – the standard defining in concrete area of professional activity of the requirement to a skill level and competence to the contents, quality and working conditions.

Development of NFQ is conducted together with the Bologna working group, thus the contents and structure of the All-European recommended frame (EFQ) is considered. The all-European recommended frame connects with each other qualification systems of various countries and acts as the instrument of transfer to make more clear qualifications outside the countries and systems. EFQs cover all range of qualifications, from basic (the 1st level, leaving school) to the advanced level (the 8th level, training of doctors). As the instrument of assistance to education during all life of EFQ covers qualifications of all levels.

According to the instruction on application of NFQ ([3]) "The national frame of qualifications defines a uniform scale of qualification levels and is a basis for system of confirmation of compliance and assignment of qualification of experts."

NSK allows:

1) describing from uniform positions of the requirement to qualification of workers and graduates when developing professional and educational standards;

2) developing estimated materials and procedures of determination of qualification of workers and graduates of all levels of professional education;

3) planning various trajectories of education conducting to obtaining concrete qualification, increase of qualification level, career development.

Now some higher education institutions of RK are performers of the international project Tempus "QUADRIGA Frame of qualifications in Central Asia: harmonization with the Bologna principles and regional cooperation"(2012-2015), its main purpose is realization of ideas of the Frame of qualifications in the countries of Central Asia and creation of the Guide on the National frame of qualifications. The main direction of implementation of the QUADRIGA project is improvement of educational systems of the partner countries by further development of the existing requirements to educational standards in certain areas on the basis of new European approaches.

Objectives of Tempus QUADRIGA project are: 1) the analysis of development of the National frame of qualifications in Kyrgyzstan, Kazakhstan and Tajikistan;

2) creation together with the Ministries of Education of the partner countries the National committees on a framework of qualifications in each of the partner countries, and also formation of a regional network of Central Asian committees on NFQ;

3) improvement of the national Regulations concerning educational standards in concrete areas on the basis of ideas of a frame of qualifications, the principles of Bologna Process and experience of higher education institutions of the EU;

4) introduction of new National standards in member countries of Central Asia on information sciences;

5) preparation of subject domain and sector descriptors for levels 6, 7 and 8 on informatics by groups of experts of National committees on a framework of qualifications.

Now with assistance of the Center of Bologna Process and the academic mobility of RK MES the methodical manuals on development of professional standards in various branches of economy of RK [6-7] are prepared. The methodical manual is developed for education and science by National Academy of Education. In this work methodological and applied aspects of formation of a branch frame of qualifications and professional standards in education and sciences are stated in detail. According [6-7] branch frame of qualifications is developed on the basis of NFQ taking into account the following principles:

- reflection of priorities of branch and accounting of business interests of the companies;
- succession and a continuity of development of qualification levels from the lowest to the highest;
- transparency of the description of qualification levels for all users;
- the description of the BFQ qualification levels through indicators of professional activity;
- the description of types of work, but not the workers who are carrying out them and workmanship them functions.

BFQs form the characteristic (descriptors) of qualification levels and subtotals opened through the main indicators of professional activity:

- knowledge;
- skills;
- personal and professional competences which are specified as follows: independence and responsibility; ability to study; communicative and social competence; professional competence.

In Kazakhstan the structure of a branch frame of qualifications includes the following elements:

1) name of branch (area of professional activity);
2) a skill level (it is specified according to NFQ);
3) qualification weight;
4) area of professional activity;
5) type of work;
6) the recommended names of positions for the allocated types of work;
7) subframes: production subframe, subframe of operational management and subframe of strategic management;
8) the description of qualification levels (NFQ descriptors for concrete qualification level are given);
9) the characteristic of qualifications on a subframes:
- the qualification subtotals allocated within concrete qualification level;
- the indicators of professional activity corresponding to each qualification subtotal the BFQ;
- indicators and descriptors, the excellent or specifying descriptors of the concrete qualification level of NFQ;
- ways of achievement of qualification of the corresponding subtotal according to NFQ.

Development of database of SQF for IT specialities

Participants of the project carried out the analysis of a condition of the national qualification frameworks; considered the European frame of qualifications with 8 levels corresponding to education levels; studied the principles of creation and introduction of the branch frameworks of qualification (BFQ) on informatics for bachelors, masters and PhD of doctors taking into account opinions of employers, in particular, the BFQ of education, the sphere of information and communication technologies, spheres of tourism and the sphere of oil and gas industry.

Today participants of the project developed the project of a database (fig. 2) for a sectoral frame of qualifications taking into account influence of professional competences on formation of the contents, the purposes and problems of disciplines of higher educational institutions.

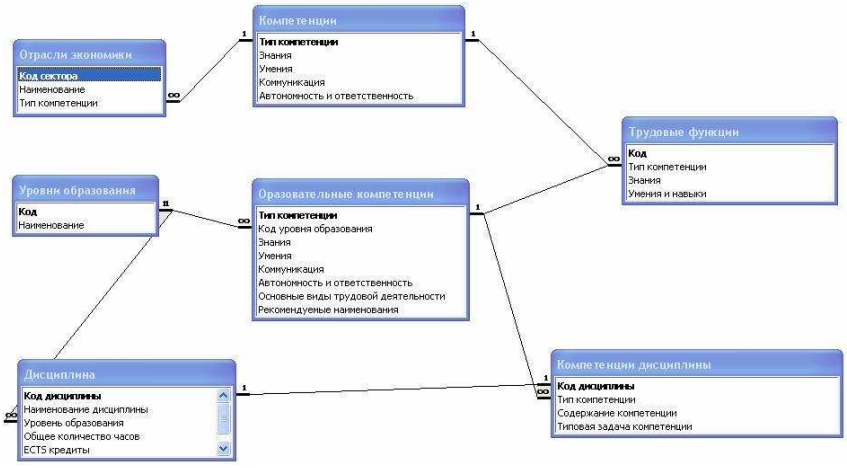

Fig. 2. Structure of database of SQF for IT specialities.

The database basis was the algorithm of development of a sectoral frame of qualifications offered by the European project coordinator Chernyshenko S.V.:

1 Transition from competences to results of training
2 Sources:
2.1 Workshop - employers - questionnaires.
2.2 Professional standards.
3 Review of base of professional standards.
4 Choice of labor functions and qualifications.
4.1 Review of statistics on graduates.
4.2 Drawing up list of positions.
4.3 Drawing up list of the companies (100).
4.4 Poll of stakeholders (employers, teachers, students, parents of students).
4.5 Forming educational standards.
4.6 Comparing competences and results of training (1: 1).
4.7 Comparing with the State obligatory standard of education (coincidence we exclude).
4.8 Ranging results of training on levels on a branch frame of qualifications.
4.9 Defining specialties and education levels.
5 Forming an educational program.
5.1 Composing the list of modules.

Conclusion

Thanks to the developed structure of a database and the information system (fig. 3), offered as the interface for the developed database, it is possible to track interrelation between the labor functions, necessary for successful work in branch of IT technologies, and educational competences of discipline; to see and estimate process of formation of professional competences of the graduate of a higher educational institution; to define the education level demanded for assimilation of necessary skills (undergraduate studies, magistracy, PhD doctoral studies), etc. In the future the team of the QUADRIGA developers will make an attempt to use this database for development of an educational program of IT specialties [8].

Fig. 3. Information system of SQF for IT specialities.

Implementation of the project will allow to provide a basis for modernization of educational programs, to promote closer cooperation and mutual understanding between universities and

employers (the enterprises, industries, public sector) that will be useful to graduates of higher educational institutions of Central Asia and will allow to strengthen the Centers of Career in higher education institutions of Central Asia for expansion of opportunities of graduates in employment and creation of high-quality employment.

Successful implementation of the QUADRIGA project will provide a basis for modernization of the higher education, and also will promote further entry of RK into world educational space.

References

1. http://www.edu.gov.kz/ru/dejatelnost/bolonskii_process/prisoedinenie_k_bolonskomu_processu/
2. The law of the Republic of Kazakhstan "About modification and additions in the Labour code of the Republic of Kazakhstan"Chapter 10-1. National system of qualifications from February 17, 2012 No. 566-IV LRK, No. 566-IV LRK.
3. The national frame of qualifications, approved by the joint order of the Ministry of Education and Science of RK from September 28, 2012, No. 444.
4. Euro-Inf Framework Standards and Accreditation Criteria for Informatics Degree Programmes, http://www.eqanie.eu/pages/quality-label/framework-standards-criteria.php (2011-09-01).
5. Computer Science Curriculum 2008: An Interim Revision of CS 2001. Report from the Interim Review Task Force, copyright ACM and IEEE, published by the Association for Computing Machinery, 2008.
6. *Seydakhmetova R.G., Kobenova G.I., Narbekova B.M., Syrymbetova L.S.* A methodical manual on development of professional standards in area of education and sciences – Astana: Center of Bologna Process and academic mobility of RK MSE, 2013. - 117 p.
7. *Seydakhmetova R. G., Pak Yu.N., Breydo I.V., Syrymbetova L.S., Narbekova B.M.* A methodical manual on development of professional standards in the sphere of power industry – Astana: Center of Bologna Process and academic mobility of RK MSE, 2013. – 95 p.
8. *Akhmed-Zaki D.Zh., Mansurova M.E., Pyrkova A.Yu., Urmashev B.A., Makashev E.P.* Sectoral frame of qualifications as basis of design of professional standards. // Materials of the 44th scientific and methodical conference "The Competence-based Focused System of an Assessment of Knowledge Vol. 1, Almaty (2014) 145-148.

Технология Создания Метрических Справочников и Конкордансов Русских Поэтических Текстов

В.Б. Барахнин[1,2], О.Ю. Кожемякина[1], А.В. Забайкин[1]

[1] Институт вычислительных технологий СО РАН, Новосибирск, Россия
[2] Новосибирский государственный университет, Новосибирск, Россия
bar@ict.nsc.ru,olgakozhemyakina@mail.ru,alexey.zabaykin@gmail.com

Аннотация. В литературоведении возникает необходимость автоматизации анализа различных уровней структуры стиха, а также автоматизированного составления на основе такого анализа метрических справочников к корпусам стихов, словарей рифм и конкордансов. Целью настоящей работы является изложение технологии автоматизации процесса создания метрических справочников и конкордансов русских поэтических текстов. Работа выполнена при частичной поддержке РФФИ (проект 13-07-00258) и президентской программы «Ведущие научные школы РФ» (грант 5006.2014.9).

Ключевые слова: автоматический анализ поэтических текстов, составление метрических справочников, составление конкордансов.

Введение

Составление метрических справочников к корпусу стихов того или иного поэта, содержащих сведения о системах стихосложения, размерах, каталектике (ритмических окончаниях стихов), строфике, метрической композиции стихотворений, а также словарей рифм и конкордансов (алфавитных перечней всех словоформ с указанием контекстов их употребления) важная задача литературоведения. Эти справочники и словари важны как для непосредственного изучения художественной техники поэта, так и в качестве основы для исследования влияния нижних уровней структуры стиха (метр, ритм, фонетика, лексика, грамматика) на высшие (речевой жанр, тематика, литературный жанр). Последняя задача является особенно актуальной, поскольку в этой области имеется целый ряд нерешенных проблем, некоторые из них сформулированы в [1]:

«Вопрос о том, связан ли метроритмический уровень текста с его тематикой, до сих пор является дискуссионным...

Методика выявления смысловой окраски ритма до сегодняшнего дня разработана недостаточно...

Вопрос этот [о тематических, образных и эмоциональных ассоциациях, связанных с теми или иными звуками — авт.] находится в стадии разработки, и пока мы не можем дать совершенно бесспорных характеристик семантики каждого звука».

Кроме того, в [1] на примере ставшей уже классической проблемы определения семантики того или иного стихотворного размера, утверждается, что методика ее решения заключается в исследовании не единичных употреблений того или иного размера, а традиций его жанрового и тематического использования, что предполагает анализ корпусов поэтических текстов. Это утверждение, очевидно, может быть отнесено и к другим проблемам исследования влияния нижних уровней структуры стиха на высшие.

Однако анализ корпусов поэтических текстов большого объема — задача чрезвычайно трудоемкая, поэтому зачастую в поле зрения исследователя попадает лишь сравнительно небольшой круг произведений поэтов-классиков, что, без сомнения, значительно снижает

полноту анализируемого материала и, следовательно, достоверность полученных результатов. Таким образом, возникает необходимость автоматизации анализа различных уровней структуры стиха, а также автоматизированного составления на основе такого анализа метрических справочников к корпусам стихов, словарей рифм и конкордансов. Это позволит освободить исследователей от рутинной работы и при этом резко расширить круг изучаемых авторов. Основные проблемы автоматизации комплексного анализа русских поэтических текстов рассмотрены нами в [2], а подходы к автоматизации процесса анализа их метрических и ритмических характеристик намечены в [3].

Целью настоящей работы является изложение технологии автоматизации процесса создания метрических справочников и конкордансов русских поэтических текстов. Как указано в [4], такой анализ должен носить комплексный характер, чтобы, проделав однажды весьма трудоемкую работу по оцифровке корпуса текстов поэта и применяя различные программы обработки поэтических текстов, мы могли получить частотный словарь, конкорданцию, словарь рифм, каталоги метрических и строфических форм и т.п.

Подходы к созданию метрических словарей и конкордансов

Первые метрические справочники к стихам русских поэтов: Пушкина и Лермонтова (работа над последним не была окончена), составлены в 1930-е годы (см. обзор [4]). Естественно, эта работа велась вручную, что требовало весьма больших трудозатрат. В конце 1960-х — начале 1970-х годов, когда компьютерные технологии обработки текстов получили достаточно широкое распространение, исследования в указанной области получили новый импульс к развитию: американскими славистами были созданы словари рифм и конкордансы к стихам Пушкина, Баратынского, Батюшкова, Тютчева (для последнего — только конкорданс), советскими литературоведами — словарь рифм Лермонтова (библиографические ссылки см. в обзоре [4]).

Разумеется, литературоведы, занимавшиеся составлением метрических словарей и конкордансов, не раскрывали детали использовавшегося ими программного обеспечения. И дело даже не в том, что для филологов этот вопрос — второстепенный. Программное обеспечение, автоматизирующее процесс составления метрических справочников, с точки зрения филолога — «ноу-хау», позволяющее получать уникальные результаты. Однако для специалистов в области компьютерной лингвистики такое программное обеспечение — непосредственный результат их научной деятельности. Так, в отделе Машинного фонда Института русского языка АН СССР был создан пакет программ UNILEX [5], предназначенный для изготовления частотных словарей, словоуказателей и конкордансов. Данный пакет был использован при создании конкорданса к стихотворениям М.Кузмина [6], при этом в статье [7] указаны его довольно существенные недостатки (отметим, что для определения количественных метрических характеристик пакет не предназначен).

Кроме конкорданса к стихотворениям М. Кузмина, за последние 25 лет были созданы Словарь языка Грибоедова [8], основную часть которого составляет алфавитно-частотный конкорданс, а также конкорданс к текстам Ломоносова [9], фактически ограничивающийся только поэтическими текстами, притом включающий лишь слова, начинающиеся на буквы А–О. Эти работы используют современные компьютерные технологии: тексты представлены на специальном языке грамматической разметки, которая основывается на «Грамматическом словаре русского языка» А.А. Зализняка, при этом предварительная грамматическая разметка корпуса выполняется при помощи программы, разработанной в компании «Яндекс», после чего проводится ручная корректировка разметки, включающая выбор вариантов разбора, снятие омонимии, разбор нераспознанных слов, исправление ошибок. В

итоге размеченный корпус текстов представляет собой базу данных, с использование которой возможно исследование различных лексических, грамматических и т.п. характеристик текстов.

Отметим важную особенностей конкордансов к текстам Грибоедова и Ломоносова: в них словарные единицы сгруппированы в гнезда лексем с указанием грамматической формы каждого словоупотребления, в то время как в конкордансах к стихам Пушкина, Баратынского, Батюшкова словарные единицы суть графемы, т. е. в одно гнездо попадают и совпадающие словоформы одной лексемы, и даже омонимы и омографы, при этом, естественно, объединение словоформ по гнездам лексем не проводилось.

Итак, для конкордансов существует автоматизированная технология их создания, в которой доля ручной работы, связанной, прежде всего, с выбором вариантов разбора и устранением омонимии, довольно велика. Эта технология сравнительно легко воспроизводима, поскольку выделения графем — задача тривиальная, а грамматический разбор слов (с указанием всех возможных вариантов, выбор из которых делается вручную) можно осуществить, например, с помощью стеммера компании «Яндекс» [10].

Однако вопросы автоматизации создания метрических справочников до сих пор исследованы очень слабо. Причины этого достаточно прозрачны: если требуемые для составления конкордансов технологии обработки текстов на уровне графем, имеющие важнейшее значение для задач информационного поиска, давно разработаны и сравнительно просты, то для фонетического анализа текстов, лежащего в основе составления метрических справочников, требуются фонетические словари, включающие, как минимум, акцентуированные (т.е. содержащие ударения) и фонетически разобранные парадигмы всех слов. Так как круг задач, требующих применения таких словарей, весьма ограничен, а алгоритмы фонетического разбора и акцентуирования неоднозначны и требуют ручной корректировки результатов, то работы в этой области ведутся не слишком активно (во всяком случае, нам неизвестны словари, удовлетворяющие сформулированным требованиям). Даже наиболее полный из известных нам сетевых фонетических словарей открытого доступа — «Словарь полного фонетического разбора» [11] — содержит только начальные формы слов, поэтому необходима генерация фонетической записи словоформ. Автоматизация этого процесса не совсем тривиальна, поскольку не существует строгих закономерностей расположения ударения в словоформах в зависимости от места его расположения в начальной форме слова.

Практически единственной работой, в которой была намечена большая программа исследований метрических, ритмических и фонетических (включая рифму) характеристик стиха, является статья [12], опирающаяся на использование системы STARLING [13]. Эта система содержит, в частности, веб-приложение для морфологического анализа [14], созданное на основе Грамматического Словаря А.А. Зализняка. Веб-приложение представляет собой морфологический анализатор, выдающий, в частности, полную акцентуированную парадигму каждого слова, имеющегося в словаре программы (к сожалению, система не позволяет генерировать парадигму произвольно заданного слова, отсутствует в ней и фонетический анализ). Рассматриваемая программа исследований характеристик стиха была частью проекта «Автоматизированный лингвостиховедческий анализ русских поэтических текстов», которым руководил С.А. Старостин, однако после его смерти в 2005 году работы по названному проекту были свернуты.

Наконец, можно отметить сайт В. Онуфриева «Рифмовед.ру» [15], посвященный стихосложению и русской рифме, который содержит, в том числе, модуль «Экспресс-анализ стихов online», позволяющий посчитать для заданного стиха количество строф, определить их тип, установить размер стихотворения, тип рифмовки и т.п. В. Онуфриев заявляет о себе как о создателе «уникальной системы классификации русских рифм», который «открыл и

объяснил новые виды русских созвучий, никем не открытые и не описанные ранее», однако точность анализа на основе его алгоритмов не слишком высока — в известном стихотворении А. Барто:

Нет, напрасно мы решили
Прокатить кота в машине:
Кот кататься не привык —
Опрокинул грузовик.

рифмовка определяется как ABCC, т. е. «решили – машине» в качестве рифмы не воспринимается, хотя это обычная неточная рифма. Особо подчеркнем: проект существует уже 13 лет, но автор не осуществил ни одной публикации в журналах, индексируемых РИНЦ, что делает практически невозможным анализ качества предложенных им алгоритмов. Теоретические же изыскания автора в области стихосложения были подвергнуты весьма резкой критике в статье [16].

Технология создания метрических справочников

При составлении метрических и строфических справочников целесообразно учитывать следующие двенадцать характеристик:

1. Количество строк, без учета пустых.
2. Метрика стихотворения.
3. Стопность.
4. Рифмовка строфики.
5. Количество мужских окончаний последних слов в стихотворных строках.
6. Количество женских окончаний последних слов в стихотворных строках.
7. Количество дактилических и др. окончаний последних слов в стихотворных строках.
8. Количество нерифмованных мужских окончаний.
9. Количество нерифмованных женских окончаний.
10. Количество нерифмованных дактилических и других окончаний.
11. Количество строк без конечных слов.
12. Тип строфической формы:

— стихотворения, состоящие из одной строфы (восемь строк или меньше);
— правильно повторяющиеся строфы;
— вольные стансы;
— парная рифмовка;
— вольная рифмовка.

Характеристики 1–4 учитываются в соответствии с метрическим справочником [17], характеристики 5-12 с конкордансом [18] (отметим, что их количественные значения взяты из [19]). Все перечисленные справочники созданы по стихам А.С.Пушкина, поэтому именно на них мы и тестировали излагаемые ниже алгоритмы.

Видимо, самым простым параметром для автоматического подсчета является количество строк (характеристика 1). Однако, и здесь есть свои подводные камни: так, в стихотворении «Когда за городом, задумчив, я брожу. . . » 17-я строка по смысловым соображениям печатается в виде двух полустрок (и, естественно, именно такой вид имеет электронная версия стихотворения), но из ритмических соображений во всех справочниках эта строка считается единой, что дает расхождение при автоматическом и при ручном подсчете строк. Выявить такие особенности графического воспроизведения стихов можно при последующем анализе рифм (полустроковая структура нарушит метрику и ритм стиха), но такая

ситуация (к счастью, весьма редко встречающаяся) потребует ручного вмешательства эксперта.

Ключевой задачей при анализе поэтических текстов является определение силлаботонических метров (характеристики 2 и 3). Для этого необходимо выделить стопу, состоящую из одного ударного слога в сильной позиции и одного или нескольких безударных. В зависимости от позиции ударения в стопе для двухсложных размеров различают ямб (ударение на четную позицию) либо хорей (ударение на нечетную позиции), для трехсложных размеров дактиль (ударение падает на 1-й слог), амфибрахий (на 2-й слог) и анапест (на 3-й слог). Для автоматического определения метрической структуры поэтического текста мы воспользовались алгоритмом, описанным в [12]. Порядок работы алгоритма предполагает построение числового вектора по следующему принципу: символом 1 обозначаются безударные слоги, 2 ударные слоги односложных слов, 3 ударные слоги, занимающие первую позицию в двусложном слове, 4 ударные слоги, занимающие вторую позицию в двусложном слове, 5 ударные слоги слов, которые длиннее двух слогов. Полученный вектор анализируется по следующим правилам:

1. Есть ли на нечетных позициях только символы 1 или 2? Если да — это ямб.

2. Есть ли на четных позициях только символы 1 или 2. Если да — это хорей.

3. Есть ли на позициях номер 2, 5, 8... только символы 1, 2 или 3, на позициях номер 3, 6, 9... — только символы 1, 2 или 4? Если да — это дактиль.

4. Есть ли на позициях номер 1, 4, 7... только символы 1, 2 или 4, на позициях номер 3, 6, 9... — только символы 1, 2 или 3? Если да — это амфибрахий.

5. Есть ли на позициях номер 1, 4, 7... только символы 1, 2 или 3, на позициях номер 2, 5, 8... — только символы 1, 2 или 4? Если да — это анапест.

6. Если 1–5 не выполнены, и отсутствует последовательность 111, это — дольник.

Характеристика 4 определяет тип рифмовки строфики. Для этого уже требуется получение фонетической информации. Фонетическая транскрипция необходима для более точного определения рифмующихся строк, нежели буквенное попарное сравнение (такие рифмы, называемые графически точными, составляют лишь небольшую долю всех рифм). Первый этап фонетической транскрипции — акцентуация — решается нами с помощью инструментария автоматической обработки текстов на естественном языке (Проект АОТ) [20], разработанного при создании системы автоматического перевода ДИАЛИНГ. Его словарь содержит порядка 3,5 миллионов акцентуированных словоформ, но, разумеется, этот словарь все равно не полон.

Для собственно фонетического анализа нами разработан модуль фонетического разбора слов, который основан на акцентуации слов с помощью последовательном (порядок важен!) применении известных правил фонетики и орфографии [21]. Следует отметить, что фонетическая транскрипция сильно зависит от ударения в слове, поэтому важно знать правильное ударение. К сожалению, это достигается не всегда из-за отмеченной выше естественной неполноты словаря ударений. Однако точность фонетического разбора в этих случаях можно повысить следующим образом. Если анализ других строф стиха (в которых проблем с акцентуацией слов не возникло) позволил нам установить его метроритмические характеристики, то на основе этих характеристик зачастую возможно установить акцентуацию слова, не входящего в словарь ударений, и провести его фонетический разбор. Вообще говоря, задача создания более или менее полной модели русской рифмы до сих пор остается не до конца исследованной, и в настоящее время нами совместно с филологами Томского государственного университета ведется ее решение, от которого во многом будет зависеть точность определения и классификации рифм.

Для определения типа рифмовки строфики при разбиении поэтического текста на четверостишья в качестве базовых вариантов проверки выделяются кольцевая, смежная, перекрестная и сквозная рифма. В случае отсутствия названных видов строф алгоритм ищет повторяющуюся структуру длиной до 16 строк. Так, в случае поэзии Пушкина максимальная длина такой структуры — 14 строк с рифмовкой ababccddehhekk (онегинская строфа).

Характеристики 5–7, отмеченные в справочнике, — количество окончаний различных типов рифм (мужской, женской и прочих) для каждого стихотворного текста. Таким образом, не учитываются различия между дактилическими и гипердактилическими окончаниями. Для определения типа рифмы в автоматическом режиме необходимо определять ударную гласную, что осуществляется с помощью упомянутого выше словаря АОТ. Известная проблема автоматизации — невозможность выбора правильного омографа при наличии разных типов омографии (падежной, межчастеречной и др.). В случае, если в конце строки стоит слово, для которого возможны разные варианты ударений, то мы не учитываем такую строку и помечаем её как некорректно определившуюся. Предполагается, что лингвист может в ручном режиме выбрать нужную форму омографа или, в случае отсутствия слова в словаре, произвести добавление слова в используемый тезаурус.

Характеристики 8–10 (количество нерифмованных окончаний последних слов в строке различных типов) определяются аналогично характеристикам 5–7 с учетом типа рифмовки. Если структура стих установлена, то найти количество нерифмованных окончаний не составляет особого труда. Более сложна ситуация, когда анализируемый поэтический текст относится к разряду свободной строфики. В этом случае привязка рифмующихся окончаний ищется в некотором диапазоне, обычно не превышающем 7.

Количество строк без конечных слов (характеристика 11) определяется посредством определения строк, выделяющихся из общей метрической структуры меньшим количеством слогов. Наконец, тип строфической формы (характеристика 12) вытекает из рифмовки строфики (характеристика 4).

Что же касается программы построения конкордансов, то алгоритм, лежащий в ее основе, достаточно тривиален и аналогичен изложенному выше алгоритму из работ [8], [9]. Основная проблема — разделение омонимов (омографов) и отнесение их к нужным гнездам лексем. В настоящее время при решении этой проблемы мы не видим альтернативы работе лингвиста (на практике — достаточно грамотного носителя языка) в ручном режиме с использованием удобного программного интерфейса

Практическая реализация алгоритма

Изложенные алгоритмы реализованы на языке программирования Python 2.7 в виде программного средства обработки стихотворного текста [22]. В процессе обработки стихотворения создается лог-файл, показывающий возникновение всех описанных выше случает неоднозначности, при этом в отдельную таблицу записываются слова, которые не были найдены в словаре ударений или у которых ударение неоднозначно. На основании этой таблицы лингвист может произвести добавление слова в используемый тезаурус или выбрать нужную форму омографа.

Тестирование алгоритма проводится, как сказано выше, на корпусе поэтических текстов А.С.Пушкина посредством сравнения полученных результатов с метрическим справочником [17] и с конкордансом [18]. В настоящее время точность определения характеристик составляет около 80 %, поэтому отладка алгоритма, наряду с повышением его точности (основанном, в частности, на экстраполяции ритмических характеристик строф стиха, в которых не возникло проблем с акцентуацией слов, на строфы с неакцентуированными

или неоднозначно акцентуированными словами), предусматривает четкое выявление сомнительных ситуаций, для которых решение будет принимать эксперт.

Заключение

Изложенная технология автоматизации процесса создания метрических справочников и конкордансов русских поэтических текстов позволяет освободить исследователей-лингвистов от рутинной работы и при этом резко расширить круг изучаемых авторов. Планируется использование этих алгоритмов для реализации программы комплексного анализа русских поэтических текстов, представленной в нашей работе [2].

Список литературы

1. Магомедова Д.М. Филологический анализ лирического стихотворения. – М.: Издательский центр «Академия», 2004.
2. Барахнин В.Б., Кожемякина О.Ю. Об автоматизации комплексного анализа русского поэтического текста / Труды Четырнадцатой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» (RCDL'2012). – Переславль-Залесский, 15-18 октября 2012 г. – С. 213-217.
3. Барахнин В.Б., Кожемякина О.Ю., Соколова О.С. Автоматизация процесса анализа метрических и ритмических характеристик русских поэтических текстов / Вестник Восточно-Казахстанского государственного технического университета им.Д.Серикбаева, сентябрь 2013, Вычислительные технологии, т. 18, совместный выпуск. Информационные и телекоммуникационные технологии. С. 248-258.
4. Баевский В.С. Справочные труды по поэзии Пушкина и его современников // Временник Пушкинской комиссии. АН СССР. Отделение литературы и языка. Пушкинская комиссия. – СПб.: Наука, 1991.– Вып. 24. – С. 65-79.
5. Аношкина Ж.Г. Лингвистический программно-источниковый пакет UNILEX+. Текстоориентированная компонента UNILEX-T // Бюллетень Машинного фонда русского языка. – 1992. – Вып. 2. – С. 3-7.
6. Гик А.В. Конкорданс к стихотворениям М. Кузмина. Т. 1-3. – М.: Языки славянской культуры, 2005-2011.
7. Бурлешин А.В. Из песеньки слов не выкинешь. . . (Рецензия на книгу Конкорданс к стихотворениям М. Кузмина. Т. 1. М., 2005) // Новое литературное обозрение. – 2006. – N 3. – С. 370-384.
8. Поляков А.Е. Словарь языка А.С.Грибоедова. – http://feb-web.ru/feb/concord/abc/
9. Поляков А.Е., Пильщиков И.А., Бергельсон М.Б. Конкорданс к текстам Ломоносова. – http://feb-web.ru/feb/lomoconc/abc/
10. Стеммер компании «Яндекс». – https://tech.yandex.ru/mystem/
11. Словарь полного фонетического разбора. – http://slovonline.ru/slovar_el_fonetic/
12. Козьмин А.В. Автоматический анализ стиха в системе Starling // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог 2006» (Бекасово, 31 мая — 4 июня 2006 г.). – М.: Издательский центр РГГУ, 2006. – С. 265-268.
13. Вавилонская Башня. Проект этимологической базы данных. Русские словари и морфология. – http://starling.rinet.ru/indexru.htm
14. Морфологический анализатор. – http://starling.rinet.ru/cgi-bin/morphque.cgi?encoding=win
15. Сайт Рифмовед.ру – http://rifmoved.ru/
16. Губайловский В. WWW-обозрение Владимира Губайловского // Новый мир. – 2002. – N 9. – С.213-216.
17. Лапшина Н.В., Романович И.К., Ярхо Б.И. Метрический справочник к стихотворениям А. С. Пушкина. – М.; Л.: Academia, 1934.
18. Shaw J.T. Pushkin: A Concordance to the Poetry: Volumes 1 and 2. – Columbus, Ohio: Slavica, 1984 / рус.пер. Шоу Дж.Т. Конкорданс к стихам А.С.Пушкина: В 2 т. – М.: Языки русской культуры, 2000.
19. Shaw J.T. Pushkin's Rhymes: A Dictionary. – Madison: Univ. of Wisconsin Press, 1974.
20. Проект АОТ. – http://nlpub.ru/AOT
21. Правила русской орфографии и пунктуации. Полный академический справочник / Под ред. В.В.Лопатина. – М: Эксмо, 2007.
22. Анализ поэтических текстов онлайн – http://poem.ict.nsc.ru/

Об Одном Подходе к Обучению Программированию

И.Н. Скопин

Институт вычислительной математики и математической геофизики СО РАН,
Новосибирский государственный университет, Новосибирск, Россия
iskopin@gmail.com

Аннотация. Анализируется сложившийся подход к преподаванию программирования, отмечаются его проблемы. Представлен альтернативный подход, связанный с отказом от использования шаблонов в преподавании и выделением двух видов деятельности в конструировании программ: составление эскиза программы, в котором не учитываются ресурсные ограничения, и отображение эскиза на конкретный вычислитель. Показаны преимущества этого подхода для развития мышления и, в частности, при обучении параллелизму.

Ключевые слова: развитие мышления, деятельностный подход к обучению, шаблоны в преподавании, проблемные задачи, параллельное программирование, модели вычислений, ресурсные ограничения.

1 Введение

Традиционное преподавание программирования исходит из предпосылки, что обучаемым следует привить навыки алгоритмического мышления на основе изучения некоторого простого языка программирования. Постепенно, по мере освоения материала расширяется круг средств языка, которые оказываются полезными для представления алгоритмов в виде программы. По своей сути это путь формирования шаблонов, комбинируемых различными доступными способами для получения алгоритмически осмысленных текстов. Он достаточно удобен в преподавании, поскольку мотивирует обучаемых к освоению нового материала.

Вместе с тем, формы представления алгоритмов в языках программирования всегда ограничивают программиста, заставляют его описывать действия из набора допустимых языком, а не излагать то, что он может себе представить на уровне осмысления задачи. В результате этого шаг за шагом программист сужает свои естественные способы оперирования данными и действиями до уровня языковых средств — формируются стереотипы мышления, которые обусловлены языком, что зачастую воспринимается как образ мышления, характерного для программирования.

Ограничительность складывающегося у программистов образа мышления проявляет себя наглядно, когда решаются задачи, предполагающие параллелизм выполнения. Это утверждение подтверждает подход к разработке параллельной программы, когда сначала строится ее последовательная версия, а потом она распараллеливается. Для поддержки такого подхода рынок предлагает специализированные системы поддержки OpenMP[1], MPI[2] и др., средства которых являются надстройками над последовательными языками.

Сопоставляя последовательное и параллельное программирование с точки зрения преподавания, стоит упомянуть, что во многом наши языки программирования наследуют свойства модели вычислений фон Неймана, постулирующей последовательное выполнение команд единственным активным элементом модели, называемым процессором. Как еще в 1975 году отмечал Дж. Бэкус [3], это обстоятельство является главным препятствием массовому переходу к программированию, опирающемуся на более развитые и выразительные модели вычислений, например, с активной памятью и с гибкой структурой организации управления.

Вывод из сказанного выше парадоксальный: нацеленность изучения программирования на развитие мышления на деле приводит к его сужению, и, возможно, единственное полезное, что получают учащиеся, это тренировка способностей приспосабливать себя к объективно существующим ограничениям. Максимально высокий уровень, который обычно в состоянии достичь программист, — комбинаторное мышление [4]. Такое мышление не способствует разработке новых методов. Они появляются не благодаря, а вопреки деятельности программистов, объективно комбинаторной по своей сути. Иными словами, методы разрабатываются за счет остатков естественных для некоторых выдающихся личностей способностей, не вытравленных рутиной программистского труда.

Это косвенно подтверждается проблемами, которые появляются у тех, кто начинает изучать параллелизм и взаимодействие автономно выполняемых процессов после освоения последовательного программирования. Всякий раз, когда приходится учитывать следствия того, что фрагменты составляемой программы будут выполняться одновременно, у обучаемого наблюдаются мысленные попытки упорядочивания процессов во времени. Как следствие, он не замечает хорошо известные ошибки, которые просто не появляются при последовательном выполнении. Ситуация похожа на то, когда программист не учитывает, что из-за округлений арифметические операции вычислителя не приводят к результатам, соответствующим хорошо изученным алгебраическим структурам. Здесь фундаментальные знания вступают в противоречие с тем, что временами случается при численных расчетах.

В последующих разделах мотивируется необходимость новых подходов к обучению параллельным вычислениям и, в частности, изучения параллельного программирования одновременно или даже раньше, чем последовательного. Это положение развивается как идея разделения конструирования программы на две составляющие его деятельности: составление эскиза, избавленного от ограничений на используемые ресурсы, и последующее отображение эскиза на реальную модель вычислений уже с учетом ограничений. Идея, иллюстрируемая рядом примеров, приводит к специальной методике эскизного программирования, которая имеет хорошие перспективы для преподавания, развивающего мышление.

Подробности изложения предлагаемого подхода можно найти в работе [5], которая рассматривает проблему применительно к раннему обучению принципам и методам параллельного программирования.

2 Мотивация раннего изучения параллельных вычислений

Наиболее эффективным методом обучения новому является создание естественной для ученика потребности. Если такой потребности нет, то нет и интереса, и, как следствие, активность учебной работы падает. Применительно к изучению параллелизма навыки последовательного программирования создают своеобразный барьер, который преодолевается, лишь тогда, когда обучаемый сталкивается с проблемой недостаточной производительности единственного процессора для решения задачи. Непосвященный в методы последовательного программирования избавлен от этого барьера, и есть основания полагать, что ему будет легче изучать параллелизм до последовательного программирования.

Приведем наглядный пример, который автор неоднократно использовал в преподавании для демонстрации параллельного решения классической задачи поиска наилучшего пути[1]

[1] Наилучший путь можно определять по-разному. Корректно считать, что это любая интегральная характеристика пути из А в Б, складывающаяся из локальных характеристик дорог. В частности, в качестве такой характеристики можно выбрать длину пути. Если предположить, что скорость перемещения для всех дорог одинакова, то этот критерий эквивалентен времени, за которое можно пройти весь путь. Представленная ниже схема не зависит от выбора критерия. Для определенности далее говорится о минимизации длины пути.

между двумя городами А и Б, связанными системой дорог. Выскажем гипотезу, что индивидуум, не обладающий навыками последовательного программирования, скорее всего, предложит алгоритм (пусть даже с вполне типичными ошибками), который можно охарактеризовать как мультиагентное решение [6]. Последовательное, а точнее, квазипараллельное решение было предложено У.-И. Далом и Ч. Хоаром для демонстрации возможностей систем с дискретными событиями языков Simula и Simula 67 в сборнике статей «Структурное программирование» [7]. Мы называем его соревнованием разбредающихся по разным дорогам агентов в скорости достижения цели [4].

Гипотетическое решение непосвященного опирается на понятие запретного города, т.е. такого, в который попадать агенту не разрешается: город считается запретным, если в нем есть или ранее побывал какой-либо агент. Первоначально все города объявляются разрешенными для посещения, и единственный существующий агент находится в городе А, а его пройденный путь пуст.

Решения описывается как поведение каждого агента, находящегося в некотором городе, по следующей схеме, в качестве двух параметров которой задаются местоположение агента (город) и длина дороги, по которой он перешел в это местоположение из предыдущего города:

1. Запомненный агентом пройденный путь пополняется его текущим местоположением, длина всего пути увеличивается на значение второго параметра
2. *Если* агент находится в городе Б, то цель достигнута. В качестве результата выдается пройденный путь;
   иначе:
   (a) Агент проверяет, является ли город запретным. Если это так, агент ликвидируется (понятно, что информация о системе в целом не теряется — другие агенты продолжают действовать).
   (b) Город, в котором стоит агент, объявляется запретным.
   (c) Порождается столько наследников агента, сколько дорог исходит из его текущего местоположения. При этом в качестве локальных данных новых агентов задается путь, пройденный родительским агентом из города А до текущего местоположения (судьба этого агента, т.е. становится он одним из экземпляров наследников или уничтожается, не принципиальна). Если из города нет других дорог, кроме той, по которой агент пришел в город, то агент ликвидируется — он зашел в тупик.
   (d) Каждый новый агент направляется на выделенную ему дорогу, пройдя которую он оказывается в состоянии 1.
3. Агент ликвидируется.

Процесс начинается порождением и активизацией единственного агента в городе А, т.е. выполнения схемы с параметрами А и 0. Вычисления завершаются, когда все агенты оказываются ликвидированными. Отметим, что достижение агентом цели не означает завершения процесса вычислений в целом. В этом случае другие агенты могут пытаться пройти по своим дорогам дальше, но это лишняя работа, поскольку судьба их — быть ликвидированными в каком-либо из запретных городов. Важно отметить, что здесь у обучаемого появляется мотив для решения проблемы завершения вычислений в общем случае параллелизма, которая в последовательном программировании просто не возникает. Это наглядный пример барьера, о котором шла речь выше. Из схемы видно, какие ошибки может допустить несведущий индивидуум, как, исправив ошибки, превратить идею в решение. Отметим ошибку схемы, существенную для методики преподавания параллелизма, которая при поверхностном рассмотрении оказывается незаметной. Она связана с одновременностью действий

агентов. Речь идет о возможных конфликтах, возникающих, когда два или более агента должны действовать в одном городе одновременно. Что при этом может происходить, и каким образом ликвидировать конфликтность, знает даже не очень квалифицированный в параллельных вычислениях специалист, и на этом вопросе мы останавливаться не будем (см., например, [8]). Какие средства синхронизации покажет преподаватель обучаемым на основе представленного примера — предмет конкретной методики. Здесь же стоит обсудить другой вопрос: как реализовать предложенное решение на реальном вычислителе. Схема действий агента не является программой для какого бы то ни было вычислителя уже потому, что она не предполагает ограничений на количество доступных процессоров, реализующих поведение. Нужен специальный ход для превращения идеи в программу. Дал и Хоар дают изящное решение, в котором действия динамически порождаемых и ликвидируемых процессов агентов регулируются так называемым управляющим списком — глобальной структурой данных, организуемой для упорядочивания вычислений специально. Это решение показывает возможность сохранения агентов в структуре реальных вычислений. Оно дает детерминированный порядок действий, выполняемых на единственном процессоре (см. [7]). Решение строится как динамическое отображение разбредающихся агентов на линейно упорядоченную структуру управляющего списка, напрямую связанного с управлением вычислениями.[2] Два хорошо известных классических решения, базирующitся на обходе в ширину и в глубину дерева всех возможных перемещений по графу, также упорядочивают вычисления, но уже за счет структуры данных, а не действий. Сравнивая три решения с точки зрения эффективности, легко прийти к выводу о том, что обход в ширину предпочтительнее. Однако при этом упускается из виду тот факт, что классические решения можно получить из агентного путем выявления всех возможных перемещений заранее, до вычислений. При обычном преподавании это обстоятельство всегда упускается, и сразу говорят о структуре дорог, пометок вершин графа и прочих понятиях, не имеющих отношения к естественному подходу несведущего программиста, представленному выше. Отметим, что задача распараллеливания классического решения для начинающего довольно трудна. Она неизбежно ведет к организации потоков и их синхронизации. По своей сути потоки есть завуалированная ипостась агентов, которой в классическом решении нет места. И именно в этом состоит для начинающего трудность распараллеливания. Возникает естественный вопрос: зачем вводить новые трудные понятия вместо рассказа о более естественных агентах? И потоковое, и агентное распараллеливание не могут обойти задачу отображения решения на конечный набор доступных процессоров. Отметим, что предложение Дала и Хоара использовать управляющий список есть ничто иное как отображение агентного решения, предполагающего неограниченность набора процессоров, на единственный процессор, т.е. обратное распараллеливанию действие. Понятно, что можно ставить задачу распараллеливания и классических решений, и агентного. Выяснение того, какая из них сложнее, выходит за рамки данной работы. Для нас важнее уже отмеченный факт сводимости агентного и классических решений друг к другу.

3 Эскиз программы и его отображение эскиза на реальный вычислитель

Представленный пример можно подвергнуть критике, указав на то, что потоковое программирование не совсем параллельное именно в связи с тем, что оно предполагает неограниченность процессорного ресурса. Это утверждение опровергается следующим образом.

[2] Решение Дала и Хоара подробно разбирается в Главе 9 «Подход к построению робастных систем взаимодействующих процессов» монографии [9]. Оно использовано как повод для обсуждения возможности использования управляющих списков при организации распределенных вычислений.

Гипотетическое решение непосвященного строится в условиях игнорирования ресурсных ограничений, а варианты доведения его до программы демонстрируют возможности построения отображения на реальный вычислитель. Эту работу можно и нужно рассматривать как самостоятельную деятельность, отличную от «неограниченного» программирования. Совмещение указанных видов деятельности для человека всегда сложнее последовательного их выполнения. Оно особенно нежелательно в учебном процессе, т.к. резко снижает его эффективность (см. [10], [11]). Именно поэтому мы провозглашаем раннее обучение параллелизму, равно как и другим общим методам, рассматривая программирование как двухэтапную схему:

– Сначала выполняется ничем не ограниченная разработка алгоритма: строится эскиз будущей программы. Степень близости полученного результата к реальной программе может быть различна, и чтобы не отвлекаться на обсуждение того, каким должен быть алгоритм без ограничений, в дальнейшем будем называть его эскизом (решения, программы или алгоритма);
– Затем строится отображение эскиза на реальный вычислитель, т.е. учет ограничений. Если иметь в виду какой-либо критерий качества, то второй этап можно считать оптимизационным, и об этом нужно говорить при преподавании программирования.

Важный аспект такого подхода — снятие на первом этапе ограничений на объем выделяемой памяти. Здесь обычная практика преподавания явно или неявно следует положению о разделении деятельностей (по-видимому, по той причине, что так проще рассказывать, а также потому, что последовательные языки не препятствуют разделению). В качестве задачи, часто используемой для проверки того, как начинающие владеют комбинационными методами, укажем на пример Гриса, в котором требуется переставить местами две последовательные части массива [12]. Дополнительное условие, требующее минимизации операций чтения и записи элементов, рассматривается лишь в качестве некоторой меры качества, а не как ограниченность времени работы процессора — этого для учебной задачи достаточно. Неограниченность памяти приводит к тривиальному решению: к переписи на новое место. Именно так поступают начинающие. А критика его есть ничто иное, как предложение построить отображение тривиального эскизного решения на вычислитель с дефицитом памяти. Приведем еще одну учебную задачу, которая с самого начала рассматривается с требованием построения параллельной программы. Она представлена в [13] для иллюстрации полезности оперирования данными, которые имеют разные одновременно существующие структуры. В данном случае речь идет о матрицах и о сравнении их диагональных структур со структурированием по строкам и столбцам. Пусть требуется преобразовать трехдиагональную матрицу к виду, в котором на главной диагонали размещаются значения среднего арифметического между самим элементом диагонали матрицы и его тремя соседями верхними, левым и диагональным верхним:

$$a_{ii} = \frac{a_{ii} + a_{i-1i} + a_{i-1i-1} + a_{ii-1}}{4} \tag{1}$$

Эта постановка задачи с самого начала ориентирует на строковое и столбцовое структурирование матрицы, поскольку апеллирует к индексам. Как следствие, программист, не очень задумываясь над постановкой, сразу же готов к реализации последовательного циклического алгоритма (по-видимому, он достаточно быстро сообразит, что нужно организовывать цикл от больших значений i к меньшим). Такое решение приводит к тому, что для реализации его на многопроцессорных архитектурах требуется специальный анализ, с помощью которого выясняется, что фактически итерации цикла не зависят друг от друга.

В то же время, данная задача очень естественно формулируется в рамках другого структурирования данных, непосредственно следующего определению трехдиагональной матрицы. В такой матрице (будем считать, что ее размерность $n \times n$) информативными являются три вектора:

- главная диагональ – вектор, длина которого равна n;
- диагональ, расположенная над главной, – вектор, длина которого равна $n-1$;
- диагональ, расположенная под главной, – вектор, длина которого равна $n-1$.

Для краткости обозначим их, соответственно, как A, A^+ и A^-.

Если определить операции конкатенации векторов $<$ вектор $> \cdot <$ вектор $>$ и отбрасывания последней (первой) компоненты $<$ вектор $>'$ ($'<$ вектор $>$)[3], а также покомпонентные арифметические операции (ниже для них использованы обозначения соответствующих скалсрных операций), то требуемое преобразование описывается как

$$A = [A_1] \cdot ((A^+ + A^- + A' + 'A)/4), \tag{2}$$

где [$<$ значение $>$]— вектор, состоящий из единственной компоненты (в данном случае используется значение A_1 первой компоненты вектора A). Сама запись задачи при втором структурировании делает ее решение естественно распараллеливаемым. Как знать, быть может, разработчики вычислительных схем, не связанные рамками фиксированного языком программирования структурирования, формулировали бы алгоритмы более пригодные для параллельных вычислений? Отметим, что разделение программирования на эскиз и отображение эскиза на вычислитель для этой простой задачи привело к оперированию со структурой, несуществующей с точки зрения языков программирования, но содержательно очень естественной. В результате построено эскизное решение, отображение которого на векторный вычислитель становится почти тривиальным. Примечательно, что для отображения его на традиционный однопроцессорный вычислитель, иначе — для представления программы в виде цикла перебора элементов матрицы, оказывается более сложным делом. Отметим, что, если векторизация требует разбиения структурных единиц на части, то отображение перестает быть совсем тривиальным, но остается достаточно простым. Эта ситуация показывает, что строить обсуждаемое отображение можно поэтапно, последовательно вводя те или иные ограничения, связанные с реальным вычислителем. Формулировка рассмотренной задачи с использованием соотношения (1) кажется более понятной, чем (2). С этим придется согласиться, но только по той причине, что мы привыкли к индексной нотации a_{ij}-ых, а оперирование составными структурными единицами считаем, если не экзотикой, то лишь специальным способом записи. Причина тому — учебные задачи, которые приходилось решать при изучении линейной алгебры, а все они, связанные с вычислениями (безразлично ручными или машинными), приводят к индексированию элементов. Если же обратиться к истокам дисциплины, то, например, задача решения системы уравнений естественно формулируется в векторной форме:

Найти такой вектор X, *что* A X = B, *где* A – *матрица, а* B – *вектор.* Векторно-матричная форма больше подходит для доказательства многих свойств операторов линейных пространств, хорошо приспособлена для объяснения ряда методов решения систем и в других случаях, когда индексы излишни. Применительно к нашей задаче было бы правильнее сформулировать ее как построение подходящего оператора в линейном пространстве.

[3] Эти операции использованы для выравнивания векторов с разным количеством компонент, которое требуется для покомпонентных действий. Быть может, правильнее было бы определить непосредственно операции выравнивания, но эта задача находится за рамками рассмотрения настоящей работы.

И тогда соотношение (1) не будет иметь преимуществ перед (2), как следствие, решение становится параллельным непосредственно. Для отображения его на реальный вычислитель остается ввести ограничение на число элементов векторов, которое допустимо для реального параллельного оперирования, и разбить диагональные структурные единицы на соответствующие части. Иллюстрацией целесообразности разделения программирования на составление эскизного алгоритма без ограничений и отображение эскиза на реальный вычислитель может служить задача построения программы, играющей в шахматы. Очень легко написать программу без ограничений, которая строит все возможные последовательности ходов, начинающиеся в исходной позиции, в виде тотального дерева, вершины которого представляют все возможные шахматные позиции. Если такое дерево построено, то можно предложить поведение компьютерного игрока как выбор одной из выигрывающих или ничейных ветвей (в принципе, любой из них), исходящих вершины текущей позиции. Отметим, что задача построения тотального дерева является естественно параллельной: для каждой позиции можно определить столько независимых потоков, сколько допустимых ходов есть в этой позиции. Поскольку тотальное дерево столь огромно, что никаких мировых вычислительных ресурсов не хватит для работы с такой «простой» программой, ее приходится рассматривать как эскизное решение, требующее отображения на реальные ресурсы и ограничивающее оперирование тотальным деревом, возможно, со снижением качества решения. Для текущей позиции вместо заранее построенного дерева при выборе очередного хода можно запрашивать построение начальных фрагментов его ветвей, урезанных на глубину фиксированного числа анализируемых ходов. В этом случае выбор хода, т.е. одного из полученных фрагментов, делается на основе какого-либо критерия предпочтения. По своей сути это урезанное отображение:

– «Бесконечное» тотальное дерево не может быть сохранено, значит, вместо этого нужно локальное построение фрагментов, фактически используемых при выборе (эта подзадача естественно параллельна);
– Выбор гарантировано выигрышной (ничейной) ветви невозможен, значит, нужен критерий, проверка которого локализуется выделенным набором фрагментов (эта подзадача также естественно параллельна: в ней выделяются параллельно выполнимые процессы вычисления значений критерия для всех фрагментов и последовательный, но некритичный по отношению к ресурсам процесс сравнения значений критерия).

Реализация представленной идеи оказывается очень выразительной при использовании средств функционального программирования, которое базируются на концепции ленивых вычислений, позволяющих задавать функции оперирования бесконечными структурами. Ленивость функции проявляется в том, что она никогда не исполняется «до конца», а выдает результаты порциями, определяемыми потребностью другой функции, вызывающей данную. В статье Дж. Хьюза [14] приведены подробности оперирования такими функциями, в частности, в применении к игровым программам, где функция построения тотального дерева «склеивается» с функцией выбора хода. Функциональное структурирование упрощает отображение эскизной программы в части оперирования бесконечной структурой, отделяя его от выбора хода. Для программы выбора хода построение отображения на ограниченные ресурсы вычислителя строится независимо от оперирования тотальным деревом, но с учетом того, что выбор нужных фрагментов дерева обеспечен.

4 Формирования метода на основе эскизов

Успех решения задачи с использованием раздельного построения эскиза и последующего его отображения уже само по себе мотивирует обучаемого к попыткам повторить опыт в подобных только что решенной задачах. По своей сути это начало формирования метода. Чтобы интерес обучаемых не угас, необходимо предложить им для эскизного решения задачу, близкую по формулировке к той, которая была только что решена. При этом есть два варианта предложения:

- задача, решение которой можно получить, модифицируя готовый эскиз и его отображение, и
- задача, для которой этот эскиз принципиально не подходит.

Первый вариант дает основание для уверенности в том, что получен новый метод, и остается только привести схему решения к независимому от задач формату, отметив при этом, какие модификации могут потребоваться. Для второго варианта нужно обсудить причины неудачи и попытаться найти пути для их преодоления. Будет ли при этом сформулирован еще один метод или нет, не столь принципиально. Важнее, что обучаемые, увидев границы применимости представленного подхода, получают представление о некорректности утверждений об универсальности методов. Эти общие положения уместно проиллюстрировать на развитии задачи о пути между городами. Представленный в разделе «Мотивация раннего изучения параллельных вычислений» эскиз подходит для обобщений, которые связаны с приписыванием весов дорогам или поиска кратчайшего пути как число пройденных городов. Такие задачи стоит обсудить для закрепления навыков использования агентного подхода, уделив им не очень много внимания. Похожая задача, для которой хочется попробовать применить готовый эскиз, — поиск всех ациклических путей из города А в Б. Первоначальный эскиз можно было бы применить, если бы как-то удалось последовательно запрещать уже пройденные пути. Однако эта особая задача, которая может вызвать затруднения у обучаемых. Проще подсказать им возможность обратного пути от Б к А, используя прежний эскиз, и изменить понятие запретных городов и поведение агента при их посещении. Дополнительно требуется решить, как будет храниться информация о найденных путях. Это одно из решений, которое связано с последующим отображением эскиза. Для новой задачи вопрос перестает быть некритичным, а потому при составлении эскиза целесообразно рассмотреть возможные варианты, один из которых утверждается, когда вводится ограничение о минимизации локальной памяти агентов и/или всеобщей для всех агентов. По результатам решения полезно сравнить решения прежней и новой задач и на этой основе показать принципы общего метода динамического программирования [15]. Отметим, что сам метод можно и естественно давать не до, а после решения задачи.

Принципиально новая похожая задача — поиск кратчайших путей ациклических для всех пар городов А и Б. Готовый агентный эскиз для ее решения подходит в тривиальном, но не выдерживающем критики варианте, когда прежнее решение запускается для всех пар А и Б, а попытке склеивания общих фрагментов путей разных агентов препятствует то, разбредающиеся агенты ничего не знают о поведении других агентов. Преодоление препятствия возможно за счет дополнительной всеобщей информации, к которой могут обращаться агенты в городе для выбора требуемого поведения. Здесь возможны разные варианты решения, которые выходят за рамки нашего обсуждения. Отметим только, что в конечном итоге может быть построена известная методика конструирования так называемых волновых алгоритмов [16]. Важно подчеркнуть, что эта методика дается обучаемым не как готовая схема, а выводится из потребности решения конкретных задач.

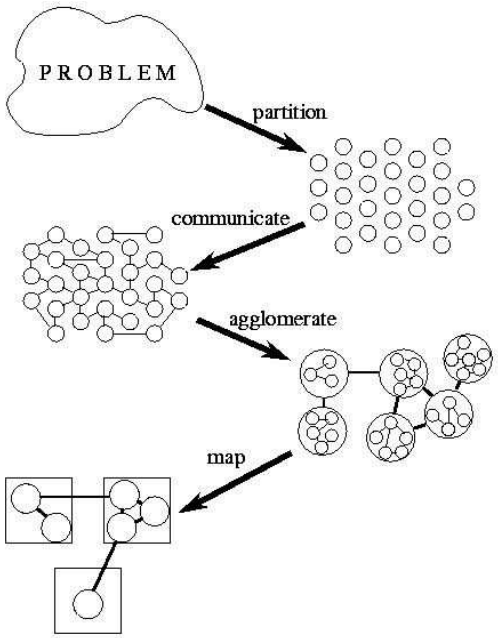

Рис. 1. Стадии методологии PCAM

Разработка эскизных решений в сочетании с их анализом, с целью обобщения в виде схем разработки параллельных программ весьма продуктивна не только в учебном плане. Достаточно общие из них могут быть предложены как специализированные методики. Примером такой методики, которая, очевидно, перекликается с агентным эскизом и может быть получена из него путем обобщения, является подход Яна Фостера (см. [17]). В этом подходе предписывается разбиение построения параллельной программы на четыре стадии: *разграничение* (partitioning), *коммуникации* (communication), *накапливание* (agglomeration) и *отображение* (mapping). Стадии методики, названной PCAM (аббревиатура по первым буквам наименований стадий), схематически представленные на рис. 1, кратко можно охарактеризовать следующим образом:

1. *Разграничение.* Выделение небольших задач, которые должны быть выполнены разрабатываемой программой, и данных, с которыми они оперируют. Вопросы, связанные с числом процессоров, на которые распределяются вычисления, игнорируются, а внимание сосредоточивается на определения возможности параллельного выполнения выделенных задач;
2. *Коммуникации.* Определяются требуемые связи между задачами, которые обусловлены необходимостью передачи данных при выполнении задач, а также соответствующие структуры данных и алгоритмы;
3. *Накапливание.* Задачи и коммуникационные структуры, определенные на первых двух стадиях, оцениваются с точки зрения требований к производительности и затрат на реализацию. При необходимости выделенные ранее задачи объединяются в более крупные с целью повышения производительности или снижения затрат на разработку;
4. *Отображение.* Каждая задача приписывается тому или иному процессору так, чтобы достигались конкурирующие цели максимизации использования процессора и затрат на передачу данных. Отображение может задаваться статически или определяется во время выполнения с помощью алгоритмов балансировки загрузки процессоров.

Последняя стадия методики явно указывает на отображение того, что построено на первых трех стадиях, т.е. на работу с эскизом конструируемой программы. То, что методы проведения работ стадий 2 – 4 не связываются с использованием формальный языков (хотя и не исключают их!) позволяют рассматривать их как эскизное программирование, последовательно добавляющее соответствующие ограничения.

Представляя свой подход, Фостер совершенно справедливо предостерегает о том, что методика PCAM не является универсальной, что она не единственная. Он демонстрирует ее применение при решении ряда задач как паттерна проектирования. Таким образом, задача обучения параллелизму решается автором в рамках традиционной методики преподавания, противоположной тому, что предлагается в настоящей работе.

5 Заключение

Представленное обсуждение, показывает, что обучение методам программирования следует связывать с общим положением о разделении программирования на две деятельности: логическое построение эскиза программы и отображение его на реальный вычислитель. Большинство примеров, рассмотренных выше, относится к раннему обучению параллельному программированию. Тому есть две причины. Во-первых, эта область демонстрирует разделение аспектов рельефнее, чем разработка последовательных программ, а во-вторых, сегодня актуальность проблем обучения методам параллельного программирования весьма высока — необходимо готовить специалистов, способных разрабатывать программные системы, которые будут корректно и эффективно использовать возможности вычислительных системах с очень большим числом процессоров. Конечно, и при разработке эскиза нужно уметь строить вычислимые эскизы, т.е. представлять их так, чтобы отображение упрощалось. Но уже сам факт разделения двух видов деятельности при программировании способствует раскрепощенному обучению, допускающему эффективную методику преподавания. Применительно к изучению параллелизму в программировании это означает, что преподаватель должен отказаться от традиционного показа шаблонных решений и сосредоточивать внимание на предложении максимально большого числа эскизов для обсуждения. Обучаемые сравнивают эскизы с разных точек зрения и в результате приходят к решению, для которого должно быть построено отображение. Иными словами, мы призываем ставить перед обучаемым проблемные задачи, которые предполагают анализ вариантов для выбора оптимального решения и имеют хорошие перспективы в отношении развития способностей разработки методов.[18] Выделение в конструировании программы деятельности построения ее эскиза с полным на то основанием можно рассматривать как программирование специального вида, который естественно называть эскизным. Эскизное программирование отличается от обычного составления программ только в одном отношении: не фиксируется используемый язык. В остальном все остается обычным: и эскизы структур данных, и операторы сохраняют свое содержательное понимание, а «ручное» отображение эскизной программы на реальный вычислитель полностью аналогично трансляции. Благодаря свободе выбора языковых средств эскизное программирование можно рассматривать как максимально высокоуровневое проектирование программ. Если же обратиться к традиционному смыслу программирования, при котором используется фиксированный язык, то такую разработку можно считать эскизным программированием, для которого компилятор реализует автоматическое отображение эскиза на реальный вычислитель. Таким образом, провозглашаемое в настоящей работе эскизное программирование можно определить как процесс, для которого компиляция остается ручной, а не автоматической работой. Это не означает, что эскизное программирование не должно быть автоматизированным. Напротив,

оно нуждается в адекватной поддержке, которая, в частности, указывала бы разработчику о возможных ошибках точно также, как и при обычной компиляции. При разработке методик раннего обучения параллельному программированию, главной проблемой, которую предстоит решить, является подбор задач, допускающих естественный для обучаемых параллелизм. Для этого следует проанализировать традиционные учебные задачи, в особенности задачи, допускающие параллельное решение, реконструируя в них эскизные составляющие и отображения на реальные вычислители. Представленные примеры могут рассматриваться как первый шаг в указанном направлении. Для традиционного обучения программирования, связывающего основы дисциплины с последовательными программами, предлагаемый подход также имеет хорошие перспективы отхода от насаждения шаблонного мышления. Здесь рекомендации все те же: ставить перед обучаемыми проблемные задачи, для которых анализ различных эскизных решений, предлагаемых учениками, должен предшествовать изложению изучаемого метода.

Список литературы

1. *The OpenMP® API specification for parallel programming.* URL: http://openmp.org/wp/ (дата обращения: 20.04.2015)
2. *Message Passing Interface Forum.* URL: http://www.mpi-forum.org/ (дата обращения: 20.04.2015)
3. J. Backus. *Can Programming Be Liberated from the von Neumann Style? A Functional Style and Its Algebra of Programs.* — Communications of the ACM. — v. 21, n. 8, August 1978. — p. 613 – 641.
4. Непейвода Н.Н. Скопин И.Н. *Основания программирования.* — Москва-Ижевск: Институт компьютерных исследований, 2003. — 868 с.
5. I.N. Skopin. *Early Learning in Parallel Programming* // In: NOVA science publishers. Series: Mathematics Research Developments. — Editor: M.S. Tarkov — 2014, pp. 219-230. ISBN: 978-1-63321-957-1.
6. M. J. Wooldridge. *An Introduction to MultiAgent Systems.* —2nd Edition, University Of Liverpool. May 2009. — 484 p. ISBN: 978-0-470-51946-2.
7. Дал У.-И., Дейкстра Э., Хоор К. *Структурное программирование.* — М.: Мир, 1975. — 246 с.
8. Хоар Ч. *Взаимодействующие последовательные процессы.* // Пер. с англ. — М.: Мир, 1989. 264 с. ISBN 5-03-001043-2
9. I.N. Skopin. *An Approach to the Construction of Robust Systems of Interacting Processes.* // In: NOVA science publishers. Series: Mathematics Research Developments. — Editor: M.S. Tarkov — 2014, pp. 199-218. ISBN: 978-1-63321-957-1.
10. Скопин И.Н. *Ролевые игры в методике обучения руководству проектной деятельностью.* // Наука и образование, 2010, № 1 (57). — с. 74 – 77.
11. Гальперин П.Я. *Четыре лекции по психологии.* // М.: Юрайт. 2000. — 112 с.
12. Грис Д. *Наука программирования.* // М.: Мир. 1984. — 416 с.
13. Скопин И.Н. *Множественное структурирование данных.* — Программирование, 2006, т.32, № 1. — с. 54 – 77.
14. J. Hughes. *Why Functional Programming Matters.* — The Computer Journal 1989, 32(2). — p. 98 – 107.
15. Беллман Р. *Динамическое программирование.* // М.: Изд-во Иностранная литература, 1960. — 400 с.
16. *Введение в распределенные алгоритмы.* // М.: МЦНМО, 2009. – 616 с.
17. I. Foster *Designing and Building Parallel Programs.* URL: http://wotug.org/parallel/books/addison-wesley/dbpp/ (дата обращения: 20.04.2015)
18. Скопин И.Н. *Проблемные задачи при изучении общих методов информатики и программирования.* — Вестник Российского университета дружбы народов. Серия: Информатизация образования. — № 4, 2011. — с. 21 – 33.

Session V. Technological Process Automation and Control

Data Processing Automation of Qeodynamic Monitoring on an Oil and Gas Field

Farida Abdoldina, Assem Berlibayeva, and Gulzada Umirova

Kazakh National Technical University after K.I. Satpayev
Almaty, Kazakhstan
{AbdoldinaFarida,Assem.Berlibayeva}@gmail.com,gulmuha@mail.ru
http://www.kazntu.kz

Abstract. The information system (IS) which automates technological processing of gravimetric monitoring raw data is presented in this article. It is carried out functional simulation of the research and production center activity which is engaged in gravimetric researches of development of an oil and gas fields to build IS data model and definition of functional part. 6 cycles gravimetric researches data conducted on one of an oil and gas fields located in southeast part of the Precaspian depression were used in the program. In work [2] is shown high efficiency of carrying out gravimetric researches for this region. The developed IS automates data and input process, ensures safety, accumulation, convenient and fast information search. IS allows to form reports, to build schedules according to parameters in inquiry of the user. Data processing automation of geodynamic monitoring on an oil and gas field allows increasing quality of processing, interpretation of gravimetric data and overall performance of the expert.

Keywords: geodynamic monitoring, gravimetric monitoring, gravitational field, oil and gas field, process model, functional modeling, information system.

1 Introduction

Widespread deployment of prospecting works, development of an oil and gas industry, increase production of hydrocarbons in the Republic of Kazakhstan under certain circumstances may create both environmental problems and bring material damage to infrastructure of the developed fields. Technogenic influences caused by development of the field may cause events of geodynamic nature due to geodynamic processes (extensive subsidence of the terrestrial surface and activate fault zones). In order to prevent manifestations of such geodynamic processes it is necessary timely implementation of high-quality geodynamic monitoring consisting in measurement gravity variations. The main objective of subsoil gravimetric monitoring is conducting long-term control of stress-strain state and fluid dynamics of geological environment, the links research of anomalous gravity variations in time with process of fields development [1]. Many companies which are engaged in carrying out gravimetric monitoring would like to have in their arsenal a certain program that would automate all processing procedures starting with the import of raw data from the file of gravimeter and finishing with maps construction of gravity variations distribution in space for various periods of time (cycles). On the other hand, the research methods implemented in finished foreign software products are somewhat differ from those which used in the CIS countries. Using of these programs does not allow to take into account peculiarities of the region of researches using software standard moreover this programs have high cost [2]. In that way, development of own software which implements fulfilled technique of processing and data presentation is practical and important task.

2 Architecture of information system

The annual volume of information obtained by geophysical surveys is tens terabyte. During the processing, interpretation and complex analysis of geological and geophysical data on computers

along with using specialized programs and automated systems there are increasingly beginning to apply information systems. A variety of information technologies allows considerably facilitate and accelerate work with geodata. Thus, information systems can be considered as effective tool to optimize storage procedures, data processing and analysis and also to provide representation of results in a readable form [2], [3]. The developed information system represents web-based application of client-server architecture which can be carried out on the Internet or an internal network of the organization. This architecture allows to centralize data storage, backup and management functions and also can support different types of clients and devices (Fig. 1). The objectives of the developed information system are automation of importing raw data, storage, accumulation, data processing and reporting.

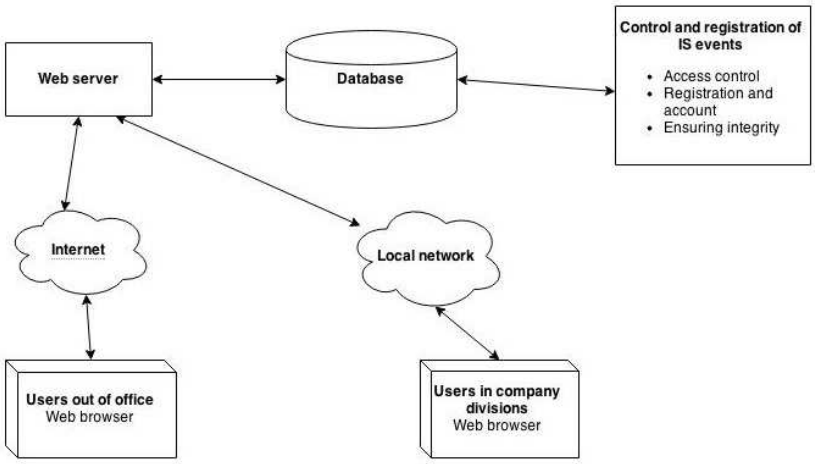

Fig. 1. Architecture of information system

3 The functional model of company activity dealing with gravimetric monitoring of development of an oil and gas fields

Creation of IS can be divided into analysis, design, coding, testing and maintenance stages. It is known that correction of mistakes made at the previous stage and it costs approximately by 10 times dearer than on current, therefore the most critical are the first stages of IS creation. It is extremely important to have effective remedies of automation of early stages IS creation [5], [16]. One of the most important stages in the design of information system is modeling of automated business processes. Qualitatively executed design of the software is an indicator of functionality and reliability [17]. For carrying out analysis and reorganization of business processes there are three methodologies of modeling: Integrated computer aided manufacturing DEFinition for Function Modeling (IDEF0), Integrated DEFinition for Process Description Capture Method (IDEF3) and data flow diagram (DFD) [16]. They are used at creation of functional model of company activity. Fig. 2 is a context diagram which represents a general description of business system and its interaction with environment. Using the IDEF0 methodology allows to create a functional model that displays structure and functions of system and information flows of material objects linking these functions.

On the context diagram the main flow of entrance information in system (Input) is allocated: area data, gravimeter data; profiles and supervision stations data; raw data obtained during

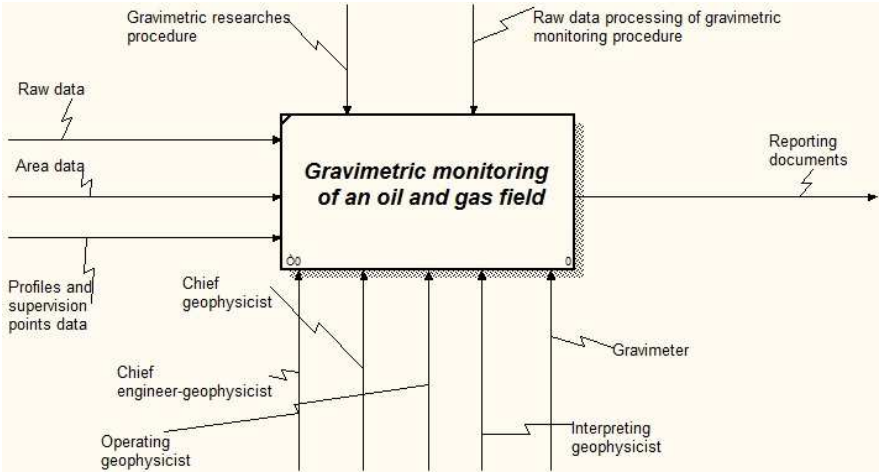

Fig. 2. A context diagram of company activity (IDEF0)

gravimetric researches. Control interface is divided into two logical blocks: gravimetric researches procedure and raw data processing of gravimetric monitoring procedure. Mechanisms are presented by chief engineer-geophysicist, chief geophysicist, operating geophysicist, interpreting geophysicist. As Output information set of reporting documents such as tables and schedules necessary for carrying out the analysis of gravitational measurements on field. Company activity can be represented as the following elements: field works, collecting and processing of raw data, preparation of reports. Available stages make a basis for creation of the first level of decomposition of a context diagram Fig. 3. Decomposition diagram designed to detail the business system functions and it turns out when splitting the context diagram into the large subsystems describing each subsystem and their interaction. Fig. 4 presents DFD model of business process

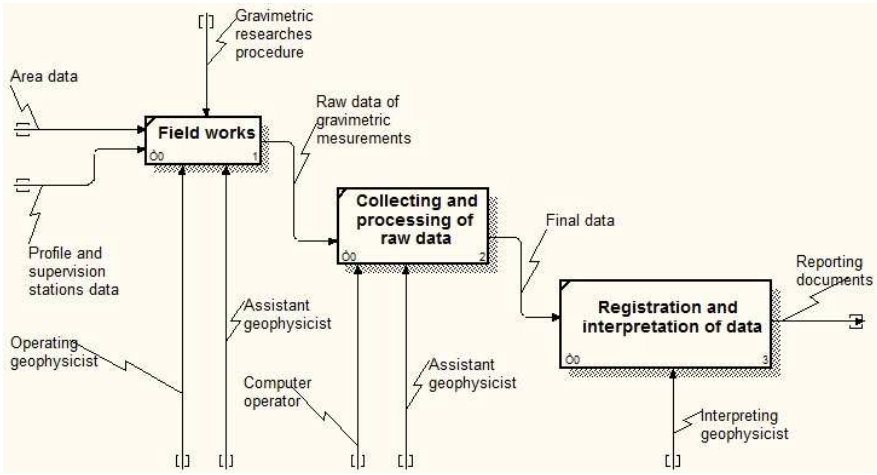

Fig. 3. The first level of decomposition of a context diagram

Data processing of gravimetric monitoring. Data flow diagram (DFD) shows how each process will transform entrance and exit data, and reveals the relations between these processes. The DFD diagram is addition to the IDEF0 model for information processing. The DFD model also represents the model-based system as network of connected works. Process Collecting and

processing of gravimetric monitoring raw data consists of the following sub runs: - raw data collection;

- information verification on parameters of gravimeter and gravitational area monitoring;
- import of raw data from files of gravimeter internal format;
- analysis of the downloaded raw data;
- calculation for each supervision station of average dG values: gravitational field, (Grav), sensitivity of gravimeter to external influences (SD), gravimeter temperatures from the level of operating temperature (Temp), humidity level (Tide), time of measurements (Time);
- calculation of gravitational field increments (dG);
- assessment of gravimetric measurements accuracy: calculation of averages dG for two gravimeters for all cycle, calculation of deviations from an average (no more than 0,010), rejection of sub-standard measurements, calculation of sum of squares deviations, calculation of accuracy on ordinary stations;
- analysis of repeated gravimetric measurements;
- creation of schedule on profiles with the changes of gravity [17].

For carrying out these processes are necessary files of the gravimeter with raw data and information on model of gravimeters, field name, supervision stations; list of profiles; list of supervision cycles; leveling data for supervision stations; data of repeated gravimetric measurements.

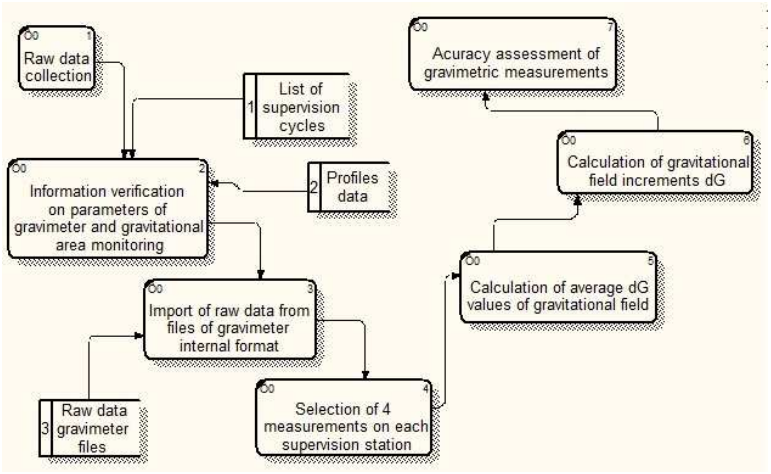

Fig. 4. Data flow diagram of Data processing of gravimetric monitoring process (DFD)

Functioning of data processing begins with function of collecting, loading of raw data which provides automated filling of system with data. Data loading are made in process of receipt them from gravimetric researches. Loaded information is a source for all other sub runs. Data selection sub run implies selection of 4 measurements on each supervision station in way the differences in values of selected measurements are the smallest. The difference between the measured variations of gravity should not exceed 0,005 mGal.

Analysis of average values of gravimetric measurements sub run calculates average values of all parameters.

The following step is Calculation of gravitational field increments (dG) sub run. Calculation algorithm of dG is based on values of gravity variations and time of measurements performance. Calculation formula is given below:

$$dG = \frac{g_o - g_{c1} - (g_{c2} - g_{c1})}{(t_{c2} - t_{c1}) * (t_o - t_{c1})} \ . \tag{1}$$

where

g_o – value of gravity variations on ordinary station;

g_{c1}– value of gravity variations on the first control station;

g_{c2} – value of gravity variations on the second control station;

t_o – value of time on ordinary station;

t_{c1} – value of time on the first control station;

t_{c2} – value of time on the second control station.

Assessment of gravimetric measurements accuracy sub run allows to analyze increment range of a gravitational field relative to strong station. Data source for this function are data received during increments calculation of gravitational field.

In Fig. 5 IDEF3 diagram for description of process Data processing of gravimetric monitoring is submitted. This diagram allows to describe interaction logic of information streams in process can be used for analysis of procedures completeness of information processing.

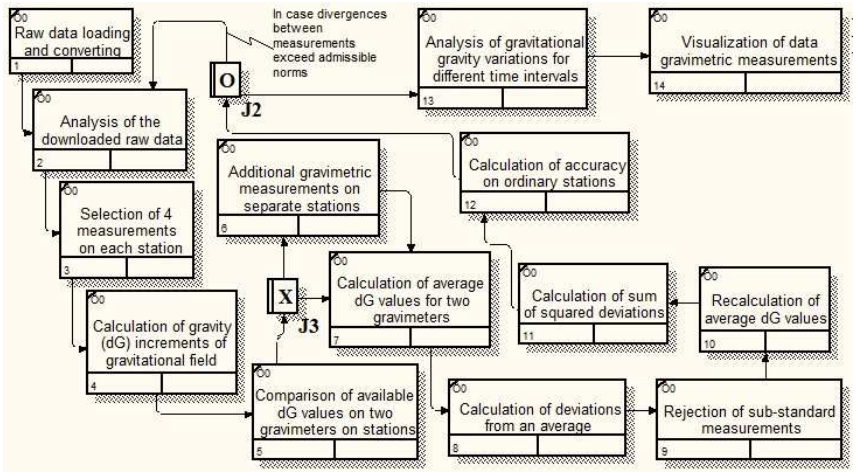

Fig. 5. IDEF3 diagram for process Data processing of gravimetric monitoring

Creation of IDEF3 model defines communications between operations. In sequence each operation is carried out only after previous one, communications between operations are displayed in the form of arrows. During performance of Calculation of accuracy on ordinary stations operation that is final level for Accuracy assessment of gravimetric measurements stage if divergence indicators between measurements exceed admissible norm, return on operation Analysis of the downloaded raw Data is carried out.

4 Model of IS data

Model of information system data is developed on the basis of functional modeling. Data model was developed consistently in several stages, as a result of which requirements to structure of data in base are defined; conceptual model of data is developed; logical model of data is created; physical model of data is developed. Fig. 6 shows ER diagram created at stage of data logical model creation.

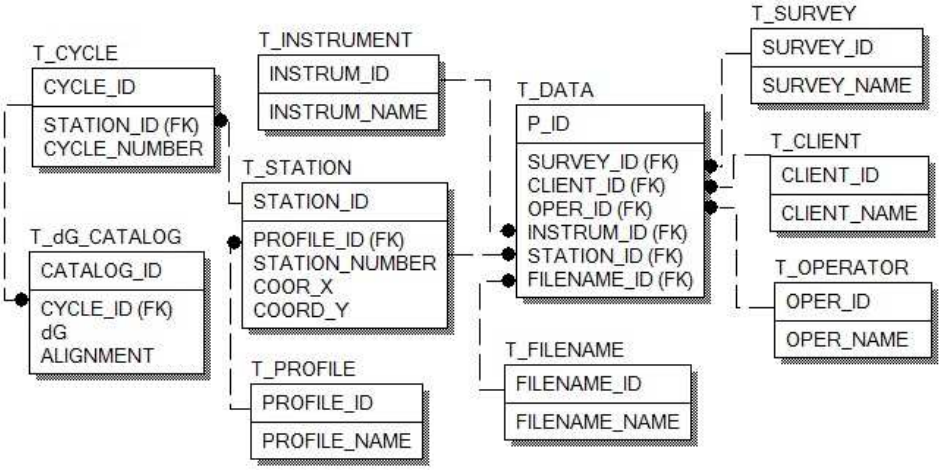

Fig. 6. Database ER diagram

Specially created loader function automates process of reading and filling of data about field name (T_ SURVEY table); client, i.e. company which is carrying out measuring works (T_ CLIENT table); models of the gravimeter (T_ INSTRUMENT table); operator who is carrying out measuring works (T_ OPERATOR table); stations on which measurements (T STATION table) and also downloaded files data (T_ FILENAME table) are performed.

The processed data of gravimetric measurements are stored in the T_ DATA table. Data about profiles on which supervision stations are located are stored in the T_ PROFILE table. Information about cycles can be found in the T CYCLE table. For amendment of dG value with the terrain data is used to leveling supervision stations which are stored in the T_ Dg_ CATALOG table.

5 IS interface and results of IS work

GravioHelper information system with convenient, intuitively clear interface is designed and developed on the basis of carefully carried out analysis of subject domain. For simplification of studying process of work with IS the user's guide is developed.

Fig. 7 illustrate system interface which visually divided into five areas they are Primary data, Average values, Accuracy assessment, Analysis and Schedule. It allows to simplify and order processing of gravimetric data. Each stage consists in data processing and results of previous stage are used in subsequent calculations.

The system of program menu consists of following points: File, Reports, Help. Through the File menu operations of opening, preservation, data change in database, press operations, data preview and export are carried out.

To start work on data processing it is necessary to download raw data from file of gravimeter (Fig. 8).

For calculation of dG it is necessary to know station concerning which all necessary calculations will be carried out. Such station is called basic. Possibility of definition by the user of a strong station does IS flexible and allows to use it for processing of gravimetric measurements on any other field (Fig. 9). Similar operations are necessary for purpose of zero station and control stations on a field. Zero station is station with which all gravimetric measurements on a field begin. On a zero station gravimeter is adjusted and given to an operating mode. Such station on which it is necessary to carry out additional measurements in connection with considerable

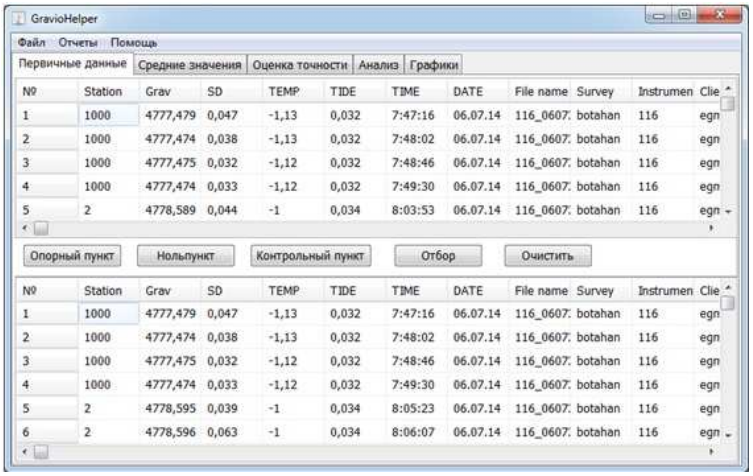

Fig. 7. IS interface for data processing of gravimetric monitoring

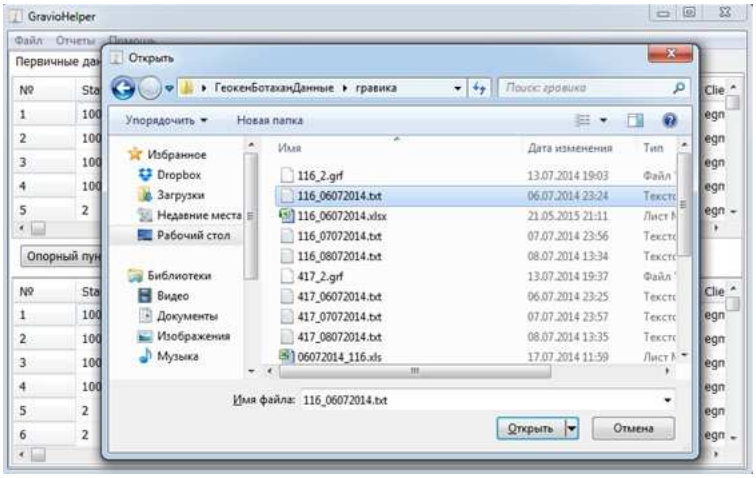

Fig. 8. Interface of gravimeter file choice

divergences of gravitational field values on two main days of supervision is called as control station.

After loading and performance of necessary settings file of raw data appears in general list of files on the Raw Data page of information system. Calculation of average Grav, SD, Temp, Tide, Time values happens on Average Values page. After determination of average sizes dG values on each station are calculated (Fig. 10). Calculation of measured increments of gravitational field by means of two gravimeters within several days on each supervision station happens on Accuracy Assessment page.

In Fig. 11 dG show values calculated by results of measurements in two main days of supervision. In the third day dG values on control stations on which additional measurements were taken are brought. All done operations are put in a basis of calculation of sum of squared deviations by method of the smallest squares. It is possible to draw a conclusion on correctness of the carried-out calculations in case of admissible value achievement of sum of squared deviations no more than 0,007 Data processing automation considerably reduced time of purpose achievement by geophysicists experts who should not recalculate manually data in case of exceeding the index of 0,007. Thus, at this stage it is already possible to draw a conclusion

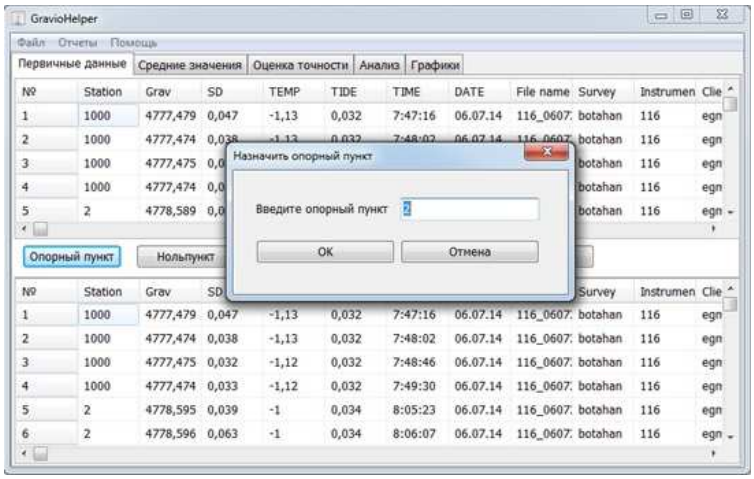

Fig. 9. Interface of strong station choice

on as far as it was important to automate data processing of gravimetric monitoring on oil and gas fields.

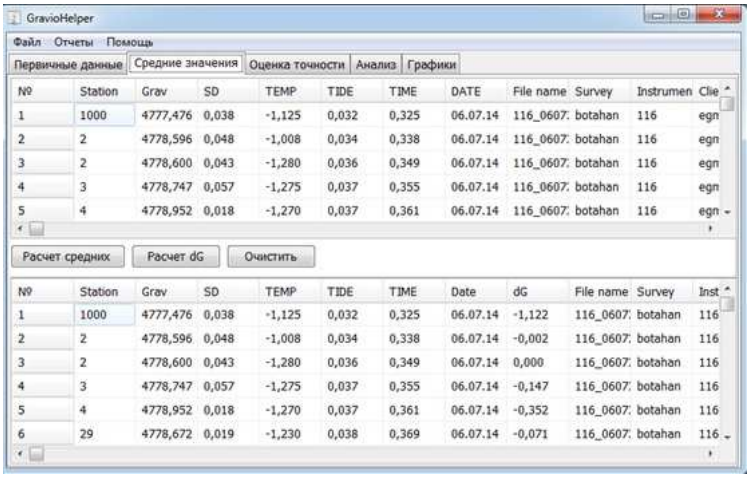

Fig. 10. Interface of dG values calculation

IS allows to accumulate and store dG values for all carried-out supervision cycles (Fig. 12). In the same window it is possible to calculate change of values of gravitational field with respect of previous supervision cycles. For graphic display of gravitational field changes in program there is a possibility of schedules creation (Fig. 13). Creation parameters of schedule are set by the user such as number of profile on which set on supervision stations, numbers of cycles on which schedule of gravitational field increment will be under construction are located.

6 Conclusion

Thus, the specialized information system is developed for gravimetric data processing automation. In system the full operation cycle of gravimetric data from data input automation, calculations of gravity values to obtaining abnormal gravity values, transformations calculation

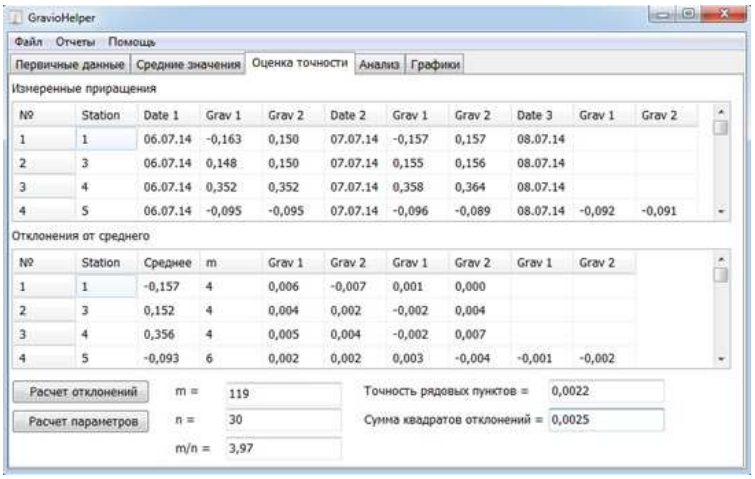

Fig. 11. Interface of accuracy assessment definition of gravimetric measurements

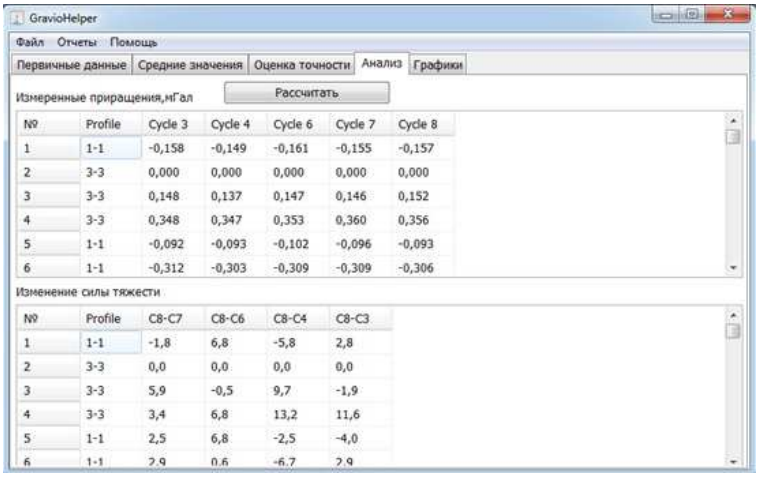

Fig. 12. Interface of repeated gravimetric measurements of gravity analysis

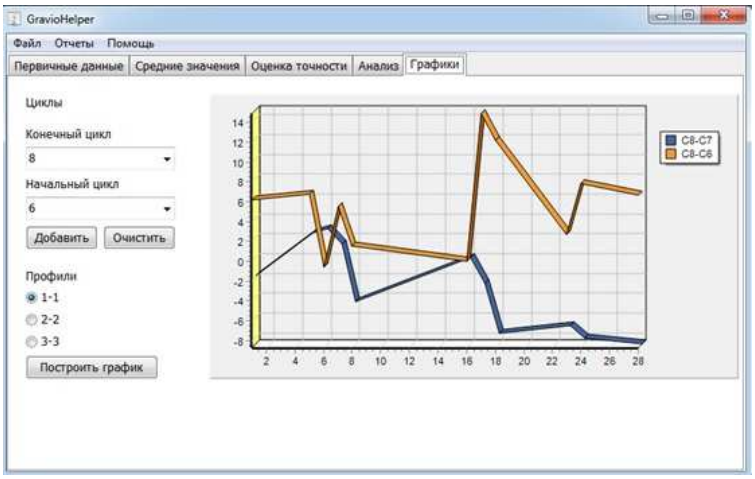

Fig. 13. Interface of schedules creation on profiles

and graphic representation of data in form of schedules is realized. Further, is planned to continue development of created information system. For visualization of data on an oil and gas field, profiles data and supervision stations, changes results of gravitational field and other necessary information it is planned to use GIS-technologies. Data processing automation of gravimetric monitoring on an oil and gas field allows to increase quality of processing, interpretation of gravimetric data and overall performance of geophysicists experts. Developed information system will be useful to companies which are engaged in prospecting works, subsoil monitoring and other works which are carried out with use of gravimetric researches [2].

References

1. Matusevich, A.: Gravitational exploration of the Precaspian depression. Livny, Mukhametov, V. (2013)
2. Naujoks, M., Kroner, C., Weise, A., Jahr, T., Krause, P. and Eisner, S.: Evaluating local hydrological modeling by temporal gravity observations and a gravimetric three-dimensional model. Geophysical Journal International, 182: 233249 (2010)
3. Skazhenik, V., Moroz, O and Vitsinsky V.: Complex of software support for computer simulation of deposits and decision of mining-geological problems // Collection of scientist papers VSB - Technical University of Ostrava, Series Mining and Geology, Vol. LI, No.2, pp. 117-122 (2005)
4. Simanov, A.: Architecture of information and analytical system of geophysical data storage and processing / Simanov, A.//KRAUNTs Bulletin. Sciences about Earth. No. 1. Release No. 7. pp. 158-162 (2006)
5. Simanov, A.: Processing and analysis of gravimetric data on basis of GIS technologies / Simanov, A. // Materials of annual scientific session of the RAHN OURO Mining institute by results of NIR in 2005 Perm, pp. 198-200 (2006)
6. Maklakov, S.: Creation of information systems with AllFusion Modeling Suite, Dialogue MIFI, Moscow (2003)
7. Maklakov, S.: CASE development tools of information systems, Dialogue MIFI, Moscow (2001)
8. Matsyashek, Leshek, A.: Analysis of requirements and design of systems. Development of information systems with use of UML: Persian English, Williams publishing house, Moscow (2002)
9. Simanov, A.: Preprocessing of gravimetric data on basis of GIS technologies / Simanov, A. // Theoretical and applied aspects of geoinformatics. Collection of scientific works. Kiyev, pp. 127-135 (2007)
10. Programs of prospecting works in the Republic of Kazakhstan for 2015-2019.

An Approach to the Development of Distributed Applications for Oil Extraction Problems

Darkhan Akhmed-Zaki, Madina Mansurova, and Bazargul Matkerim

Al-Farabi Kazakh National University,
al-Farabi av., 71, Almaty, Kazakhstan
http://www.kaznu.kz

Abstract. Paper presents an attempt to develop and deploy a functioning MDA (Model-Driven Architecture) model of a distributed application for oil extraction problem. Main focus is the problem of modeling distributed computing processes using MapReduce paradigm in a visual notation and automatically generating executable code using resulting diagrams. Article reflects efforts on creating platform independent model of process execution, transforming it into the platform specific model, and, finally, automatically generating application code. Novelty of the research include platform independent model of the classic hydrodynamics problem, equivalent Hadoop based platform specific model and testing results that confirm feasibility of the research.

Keywords: MDA technology, MapReduce Hadoop, oil extraction problems, distributed scientific computing.

Introduction

At present, an effective large-scale computing problem of oil and gas industry solution is related to the use of high performance computing technologies. According to [1] "among 500 most powerful supercomputers of the world, their use in geophysics by the companies of oil and gas service for prospecting and exploratory development of deposits holds the third place". Thus, the choice of the corresponding parallel hardware and parallel models of computing depending on volumes of scientific computing and specifics of problems is an actual problem.

In many cases, problems of oil and gas industry come to modeling of oil reservoir. A traditional process of developing applications for an oil reservoir consists of several phases:

- Investigation on the peculiarities of an oil reservoir;
- Mathematical modeling of physical processes;
- Numerical modeling of processes;
- Development of software;
- Validation and verification of software;
- Testing of programs.

When developing complex high performance applications, there may appear errors at different stages due to the dependence on hardware, ambiguity in interpretation of models, complexity of programming [2]. The transfer of a programming code from one parallel architecture onto another one, modification of software, documentation, protection of the established software from external effects are only part of problems which necessary to solve, when implementing and maintaining such applications.

Thus, to eliminate the above mentioned issues in the course of designing, development and maintenance of applications in oil and gas industry, we propose to use special methodology of software engineering. Investigations on modern approaches to designing and development of software showed that, when developing applications for high performance scientific computing

(HPSC), MDD methodology gives good results [3]. On this basis MDD methodology is chosen by authors to design and development of HPSC applications for the oil and gas industry. The technology MDA is widely used for development of applications in different fields: financial, commercial, state, educational [7]. Besides, MDA is used for development of applications in the scientific sphere. The use of MDA technology for developing applications for high performance computing in the oil and gas industry is described below.

The problem statement

Let us consider a hypercube in anisotropic elastic porous medium $\Omega = [0,T] \times K\{0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}$.

Let equation (1) describes the fluid dynamics in hypercube Ω under initial conditions (2) and boundary conditions:

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x}(\phi(x,y,z)\frac{\partial P}{\partial x}) + \frac{\partial}{\partial y}(\phi(x,y,z)\frac{\partial P}{\partial y}) + \frac{\partial}{\partial z}(\phi(x,y,z)\frac{\partial P}{\partial z}). \quad (1)$$

$$P(0,x,y,z) = \varphi(0,x,y,z). \quad (2)$$

$$\frac{\partial P}{\partial n}|_{\Gamma} = 0. \quad (3)$$

Here under conditions (3) is the surface of cube Ω. In equation (1) the solution function *P(t, x, y, z)* is seam pressure in point *(x, y, z)* at the moment t; ϕ*(x, y, z)* is diffusion coefficient in the reservoir; *f(x, y, z)* is density of sources. To solve (1)-(3) Jacobi's numerical method was used. For problem (1)-(3) a parallel algorithm of solution was realized using MPJ technology presented in [2].

The technologies for design and development distributed application

The concepts of MDA technology

Model Driven Development (MDD) is a methodology according to which, when creating software, models become the main objects of development. The code and other objects are generated further from models. Here, a model is a formal specification of the function, structure and actions of the system [1]. According to MDD methods, all stages of designing and development of applications are based on models, therefore, this model is called model-driven. At present, an MDD method in the field of information technologies is one more level of abstraction in the object-oriented approach, when developing software. The increase in the level of abstraction facilitates the process of designing and development of application.

At the beginning of 2000, OMG consortium [2] initiated a conceptual infrastructure of MDA and its main principles for MDD methods as well as standard languages of specifications for description and transformation of models.

Figure 1 presents generic development cycle an MDA application. It consists of three stages: analysis, design and code implementation. At the first stage, special attention is paid to development of PIM model, at the second stage to PSM and the third stage to writing a programming code.

Development of models means not only their understanding, description but also presentation in the form in which they can be processed on the computer as well as the possibility to transform

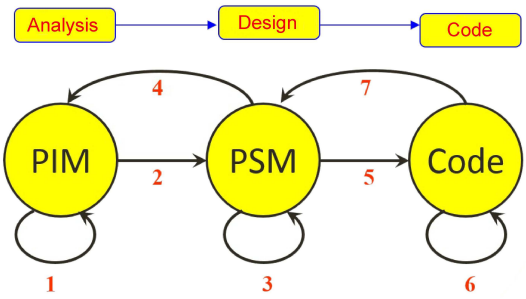

Fig. 1. Figure illustrates MDA application development cycle [3].

PIM model to PSM model, generation of a programming code from PSM model. In the figure, near the arrows connecting the elements, there are numbers denoting subsequent steps of the system development cycle. An initial prototype of the model may contain inaccuracies or discrepancies, therefore, the subsequent prototypes are created by validation and introduction of changes. Only when the created model undergoes testing and corresponds to the requirements this model can be transformed to the next model. In the cases when it is necessary to introduce changes to the functioning system, for example, to start it with the help of other technology, models PIM, PSM should be reconsidered (Figure 1, steps 4, 7). Then, taking into account the changes, it is necessary to create new PSM model (for example, steps 2, 5) and generate the code.

To create MDA models, the Unified Modeling Language (UML) can be used. In this case for the development of MDA applications the following diagrams of UML language: use case diagrams, class diagrams, state diagrams, diagrams of interaction are used.

Using MDA technology for design distributed application

Figure 2 shows the process of designing and developing applications for HPC using UML activity diagram. It is seen in the figure that the general process is divided into 5 stages. Each stage is realized by a specialist indicated on the left. The result of the specialist's activity is the corresponding MDA model shown on the right.

Designing and development of an application for HPC are carried out by a group of specialists with the help of the approach with passing on the baton (Figure 3). From the experience of high performance scientific computing application development for oil-gas industry we found that a concerted cooperation of specialists from different fields is important. Therefore, we recommend to organize "Relay race of specialists"in high performance scientific computing application development. Cooperation of specialists can best be described with the help of the approach of passing a baton.

A specialist in the oil-gas industry defines the subject of research and makes the statement of the problem. A specialist to whom the baton was passed creates a mathematical model of the problem. Then the baton is handled on to a specialist in the field of numerical methods who finds a corresponding explicit or implicit numerical method. The following stage is the work of a high performance scientific computing application designer. In the implementation stage the baton is passed to a high performance computing specialist, who is responsible for a final programming product. In the process of studying specifics of HPSC computations we developed four high per-formance scientific computing components with the help of which different HPSC applications can be constructed [2]. The developed components are divided and named depending on the peculiarities of HPSC application. It is of no importance for solution of what scientific computing problems this application will be created, its model can be

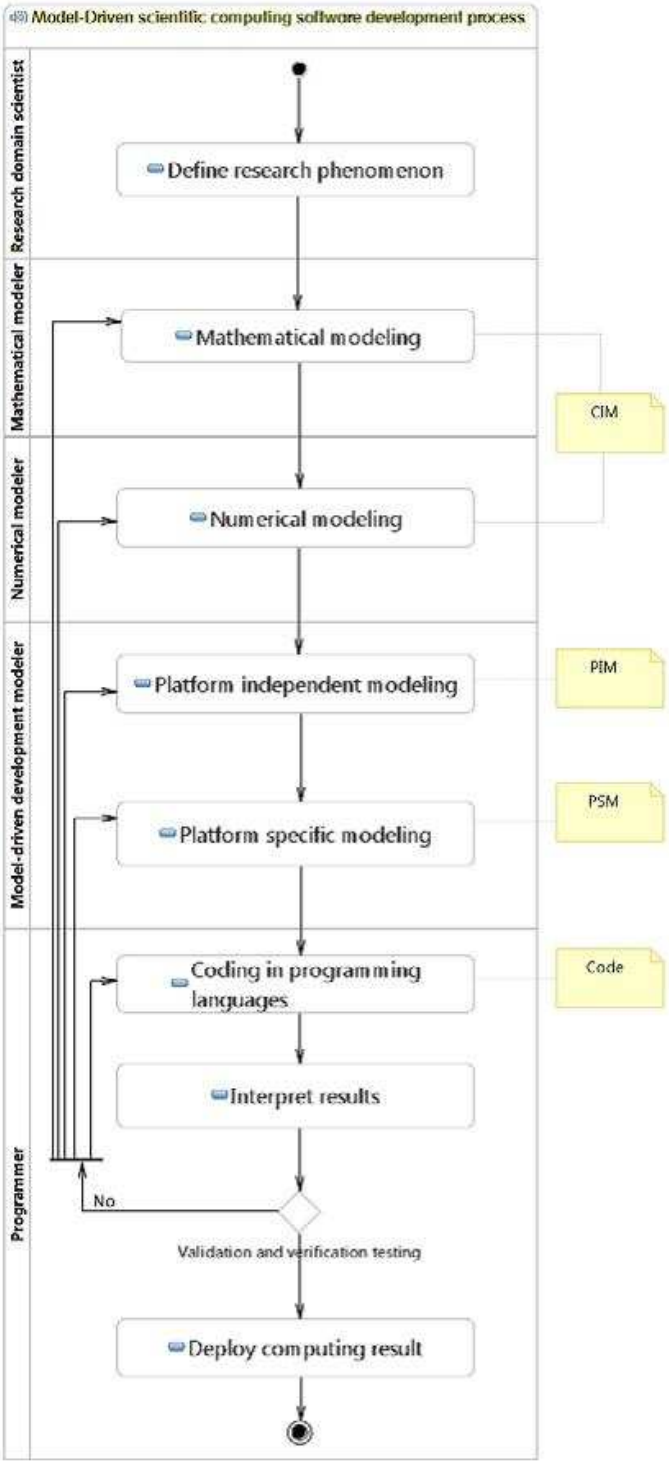

Fig. 2. MDD process of application development

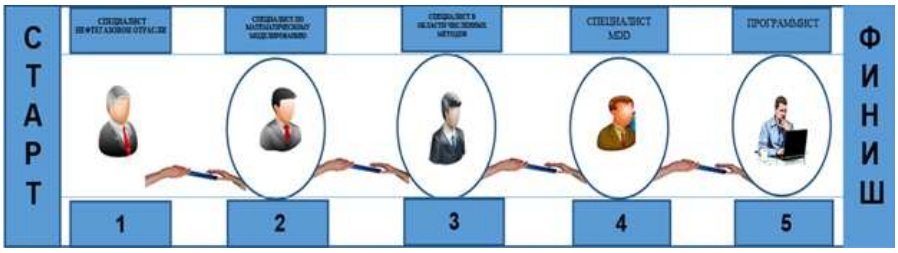

Fig. 3. Relay race of specialists in application development.

constructed from these components. Every time at the stage of designing an application for a computing problem we have to determine general input and output parameters, class of equations, explicit or inexplicit methods of computations and instruments (tools) for performing parallel computations. Therefore, we presented these 4 invariable independent parts in the form of 4 basic components. They are: an input-output component. InOutPut, a component of equations of problem statement SciEquations, a component of numerical methods NumerMethods and a component of organization of a high performance computing medium PEOrganize. At the previous stage of investigations, using these components we carried out designing of an application for scientific computations with the help of MPI technology. In this work, we set up the task to develop an application with the help of these basic components using the technology MapReduce Hadoop. In the model of the earlier created HPSC MPI application is compared component-by-component with the MapReduce application under study, the difference is found to be only in component PEOrganize which is responsible for organization of high performance computations. The remaining components are the same. Component PEOrganize is developed taking into account the peculiarities of MapReduce technology, then the relation of this component with other components is realized with the help of independent classes entering into these components. Every of the presented component consists of several classes. Description of the components as applied to MapReduce application is presented below:

1. Component PEOrganize is used for creation of the topology on MapReduce Hadoop platform.
2. Component NumerMethods is used for determination of a numerical model with different types of grid and numerical methods.
3. Component SciEquations used for determination of a mathematical model with the number of final differential equations and conditions for these equations.
4. Component InOutPut consist of classes of reading from the file and writing into the file with the help of which input and output data of HPSC application are prescribed (set).

As is shown in Figure 4, in the process designing and development of applications there takes place transformation of models starting from the upper level to the lower level.

As it was mentioned above, MDD - specialist receivers a computationally independent model CIM from the specialist in numerical methods. In case of solution of problem (1)-(1) model CIM contains the algorithm of a numerical salutation of the problem by the explicit method. In his turn, MDD-specialist creates an independent on the platform and the programming language PIM model for the given numerical CIM model using HPSC components. Model CIM can be described by the components of input-output-InOutPut, the component of equation of a scientific computing problem - SciEquations, the component of numerical methods - NumerMethods.

But in CIM model there is no information for the component of organization of high performance computational (computations) medium - PEOrganize as the medium of development

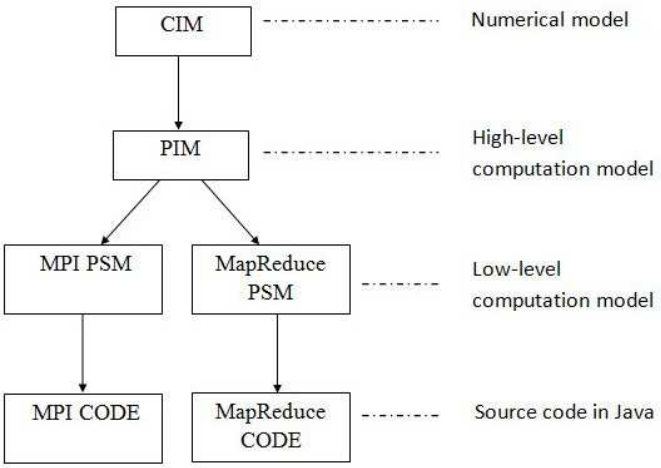

Fig. 4. MDA modeling.

in the computational models is not considered. The work resulted in the development of MDA model and realization using the MapReduce Hadoop technology.

The distributed computation algorithm based on MapReduce Hadoop technology

According to MapReduce paradigm, realization of the algorithm for iteration processes is a series of MapReduce tasks [15-17]. First, original domain is divided into sub-domains (Figure 5). Every sub-domain consists of three main parts: ghost slab, boundary slab and interior slab. Data transformations (defined by the algorithm) can proceed independently only in interior slab. Boundary slab of the sub-domain that is being computed requires boundary slab values of its neighbors and those are stored in the ghost slab. In other words, ghost slab stores copies of neighbors' boundary slab values.

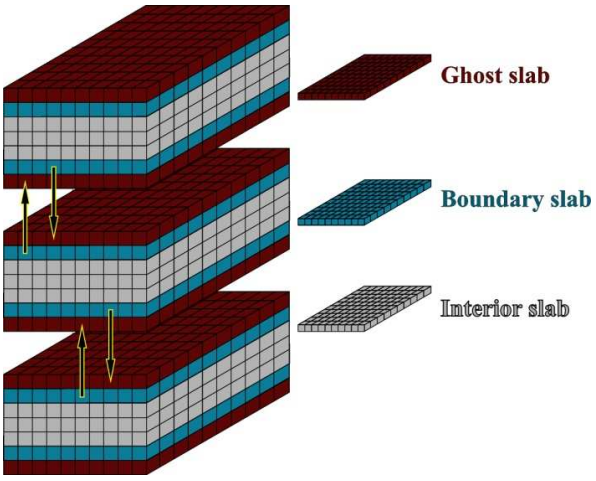

Fig. 5. The hypercube is divided into number of computation sub-domains. Every sub-domain is assigned a reducer.

Algorithm of numerical solution of the problem (1)-(3) with the help of MapReduce Hadoop technology consists of two stages: the stage of initialization at which MapReduce work of the first

level is performed only once and the iteration stage at which a cycle of MapReduce works of the second level is performed. Mapper of the first level loads data from the file system HDFS. Then, Mapper distributes the data between Reducer processes on slabs, thus realizing 1D decomposition of the data.

Reducer, in its turn, performs computations, duplications boundary slabs into the ghost slabs of the neighbors and stores the obtained results. The data used by Reducer for computations are divided into two kinds: local data, i.e. the data which refer to the interior slab and shared boundary data (boundary slab). Reducer enters local computed data directly into a local file system and enter the shared boundary data into the output file of the distributed file system HDFS, which will be an input file for Mapper of the second level at the next iteration. At each iteration Mapper of the second level distributes the updated boundary data among Reducers, thus providing the exchange of boundary values between slabs. The fluid of data corresponding to the description is presented in Figure 6.

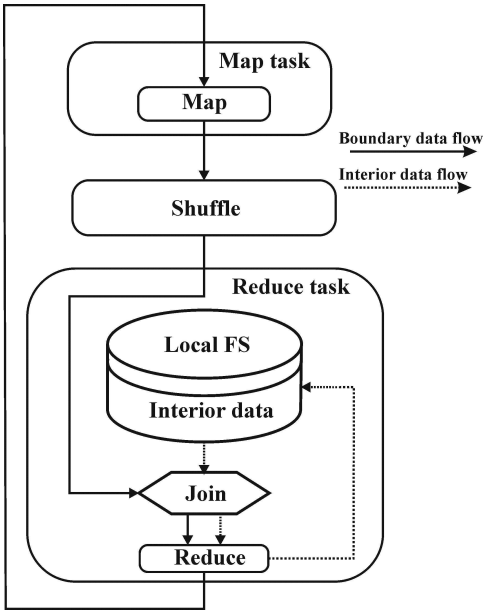

Fig. 6. The hypercube is divided into number of computation sub-domains. Every sub-domain is assigned a reducer.

The distributed algorithm consists of two stages:

- The stage of initialization;
- The iteration stage.

The stage of initialization is a MapReduce task Initial in which there takes place initialization and writing of files necessary for computations in the process of iterations. The iteration stage is a MapReduce task Iterations. At each iteration in Mapper, points of the field with the same keys, i.e. numbers of subcubes, are grouped. The input data of Mapper are the output data of Reducer. In Reducer, the main computations are performed according to the formulae of the explicit method. Then, writing of the interior parts of files into the local file system and transfer of values of boundary slabs to the output of Reducer Iterations are performed.

Solution design for the problem

PIM model

In our case, computations are performed in MapReduce Hadoop medium. The algorithm with the use of MapReduce consists of the stage of initialization and iteration stage, a separate MapReduce work being fulfilled at each iteration.

Computations are performed on Hadoop platform, a MapReduce problem receives a cube of data, Mappers perform 1D decomposition, and each Reducer receives its block of data and performs computations. After computations are completed, boundary data are entered into a distributed file system HDFS, the values of inside points are written into a local file system. Then a new cycle begins. The process continues until the condition is satisfied. Thus, we have developed a PIM model for HPSC applications for problem (1)-(3) with the help of UML diagram of classes (Figure 7) indicating relations between the classes. On account of retrieving calling methods of each other, the classes of components are in associative relations.

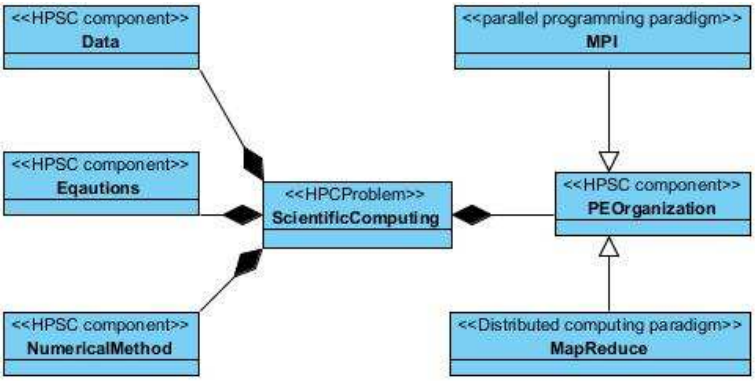

Fig. 7. Figure PIM model.

PSM model

Models of transformation of PIM to PSM can be classified by several categories [10]: improving the quality of transformation, with perfection of the development, with refining, with specialization, translation, abstraction, generalization and forms of designing. In our case, transition from PIM model to MapReduce Java PSM model refers to the category with refining. Refining means redetermination in the course of transition from CIM to PIM, transition from PIM to PSM. Refining can be added at one level of abstraction transformation of PIM model to MapReduce Java PSM model is transformation of UML-diagram of classes Java to the diagram of classes Java with addition of MapReduce specification to PSM. When transformation PIM model to MapReduce Java PSM model, multiple succession, associations of classes and qualified associations must be removed. In PSM model, the relations between components shown in Figure 8 are preserved, but specification of the programming language Java is added.

An automate transition from MapReduce Java PSM to Java code is realized with the help of generator Acceleo [11]. Acceleo is a pragmatic realization of Object Management Group (OMG) of MOF model. Acceleo UML2 for Java is a code generator based on Acceleo 3.2. This generator supports creation of the initial code Java for classes and interfaces. Following figure presents test run results:

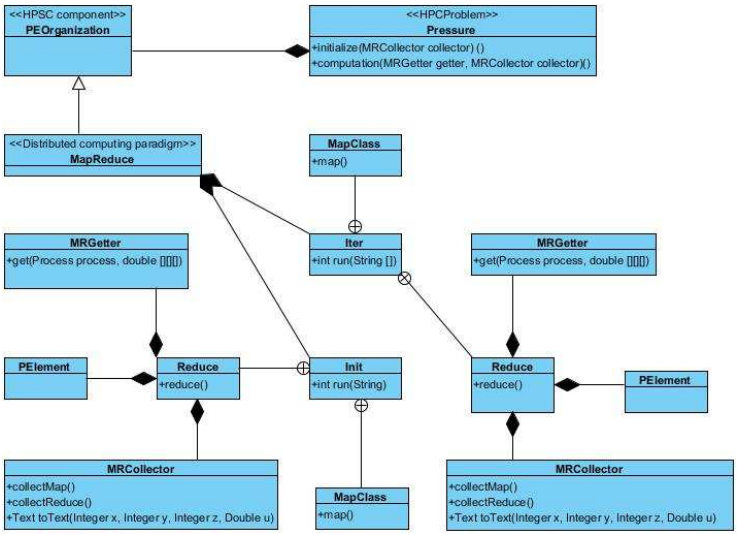

Fig. 8. Figure PIM model.

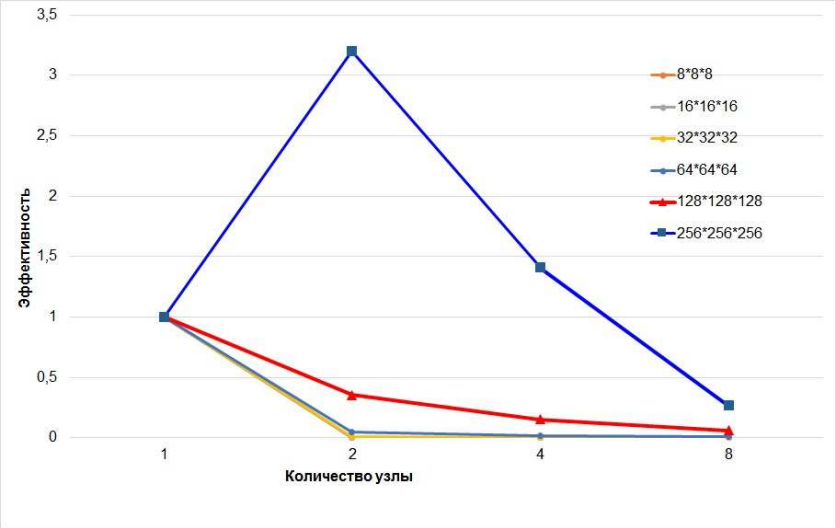

Fig. 9. Efficiency.

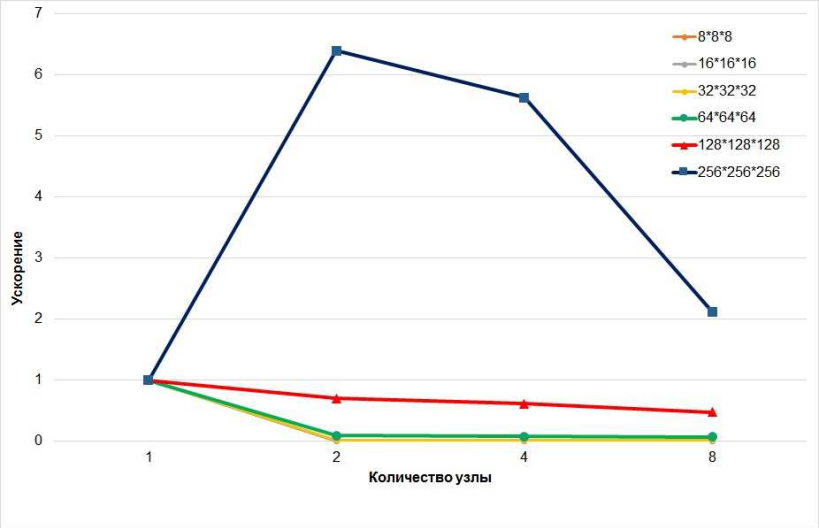

Fig. 10. Speedup.

Conclusion and future research

The aim of this direction of investigations is the use of MDD methods for development of applications for HPC in the field of oil and gas production. For the work, MDA standard is chosen as MDD methodology. The process of designing and developing a high performance application is described on the example of MDA modeling. The method of passing on a baton between different specialists of oil and gas industry, the close interaction of which can facilitate the work on creation of complex applications for oil and gas industry, is shown. The investigation results show the prospects of using MDD methodology for solution of complex recourse intensive problems.

The experimental results allow to conclude that the distributed application works well and with the increase in the volume of the data being processed the performance of Hadoop implementation increases. HPSC applications can be designed and developed with the help of the proposed MDA model and its basic components. This approach will possible one of the ways to do distributed scientific computing on high performance heterogeneous systems.

Acknowledgments

Presented research was funded under Kazakhstan government research grant.

References

1. Tokarev M.Y., KurinE.A., Sinitsyn M.N. Supercomputer technologies in science, education and industry. Edited by VA Sadovnichy, M .: MGU, 2012.
2. Lugato J., Bruel M., Ober I. Model-Driven Engineering for High Performance Computing Applications // Proc. Modeling Simulation and Optimization Focus on Applications.-2010. -P. 303-308.
3. B. Matkerim, D. Akhmed-Zaki, M. Barata, Development High Performance Scientific Computing Application Using Model-Driven Architecture, Applied Mathematical Sciences, Vol. 7, 2013, no. 100, 4961-4974.
4. Daniluk. Visual modeling for scientific software architecture design. A Practical approach // Computer Physics Communications. -2012. -No.183. - P. 213-230.
5. Ober MDE4HPC: An Approach for Using Model-Driven Engineering in High-Performance Computing, SDL 2011, LNCS7083, pp, 247-261, 2011.
6. Scott W. Ambler. The Elements of UML 2.0 Style, Cambridge University Press, 2005. 7. MDA success story. URL: http://www.omg.org/mda/products-success.htm.

Technologies of Heterogeneous Programming Systems Integration in the Informational Computing Environment of Mathematical Modeling and Data Analysis *

Igor V. Bychkov, Gennady M. Ruzhnikov, Roman K. Fedorov, Alexander S. Shumilov, Ivan A. Sidorov, and Vadim P. Potapov

Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of Russian Academy of Sciences,
Irkutsk, Russia,
Institute of Computational Technologies of Siberian Branch of Russian Academy of Sciences, Kemerovo, Russia
{bychkov,rugnikov,fedorov}@icc.ru,shumsan1011@gmail.com,yvan.sidorov@gmail.com,kembict@gmail.com
http://idstu.irk.ru

Abstract. Usually the solution of complex high-technology tasks requires usage of distributed programming systems realized on completely different programming platforms. The technologies of WPS-services creation were developed in ISDCT SB RAS. The proposed technologies allow to simplify service creation in cloud environment, supercomputers and to combine distributed program systems into single application.

Keywords: OGC, WPS, SOA, REST, Cloud computing, Computing cluster.

1 Introduction

Usually the solution of complex high-technology tasks requires usage of distributed programming systems realized on completely different programming platforms. For instance, the problem of modeling of the atmosphere pollution spread in inhabited localities is one of those problems. The solution of the problem can be divided in two parts: monitoring of pollution sources and modeling of the a pollution spread. The realization of pollution sources monitoring is usually based on geo information systems. The modeling of the pollution spread requires significant computing resources and can be realized using supercomputers. Therefore the program solutions of the mentioned parts of the problem can be realized by various groups, on various operating systems, using various development studios etc. Nowadays Service-Oriented Architecture (SOA) [1] is actively developing using standards of Open Geospatial Consortium (OGC). One of the most promising standards is Web Processing Service (WPS) [2], which unifies the use of the Internet services that deal with the spatial data processing. The mentioned program systems interfaces can be realized using this standard. It allows to invoke them from any part of the Internet. SOA is widely used and most promising interface standard for software development on the base of Cloud computing [3]. Each program has its own set of system requirements for the software and hardware environment. The use of cloud computing allows us to provide the necessary environment and solve the problem of compatibility and computational resources for services. Therefore problems of service creation and publication in cloud environment and the problem of composition of traditional servers and supercomputers arise. The informational computing environment of mathematical modeling and data analysis [4] is being developed in the ISDCT SB RAS. The environment is based on web technologies and provides the subsystem

* The work was supported by RFBR 13-07-12080 (ofi_m), program of Presidium of RAS "Fundamental problems of mathematical modeling".

of data storage and the set of data analysis WPS-services for user. Any WPS-based service can be integrated in the system by registering it in the environment's services catalog.

Certain technologies, based on existing programming systems and that make the implementation of WPS-services easier, were developed:

– the technology of WPS-services creation that is based on cloud computing technologies, that provides virtual machines in cloud infrastructure with the set of preinstalled software that allows easy implementation of different programming systems as WPS-services;
– the technology of WPS-service creation based on computing cluster that allows to access programming systems of computing clusters.

2 Technology of Creation of WPS Services Using Cloud Computing

Technology of creation of WPS services using cloud computing consists of following steps:

1. Providing a virtual machine in the cloud.
2. Software installation and configuration, data duplication.
3. Setting up the application software in the form of WPS service. There are several variants, depending on the environment of software implementation: description of the template for running software on the console; implementation in the form of libraries for Zoo-project or 52°North WPS.
4. Registering WPS service in the catalog.

Consider more detailed providing a virtual machine in the cloud. The architecture of the components to create virtual machines is shown in (Fig. 1).

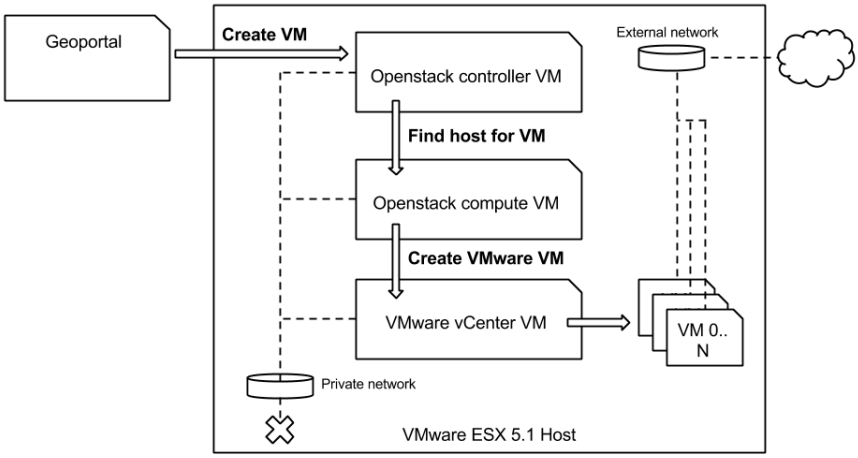

Fig. 1. The architecture of component for virtual machine creation.

The virtual infrastructure of Geoportal is located on a physical server, which is under control of VMware ESXi, which provides low-level API for managing objects of the virtual infrastructure. In this physical server, in addition to established client virtual machines there are following virtual service machines:

1. VMware vCenter -– virtual machine based on Windows Server 2008 R2, providing API to manage physical VMware server. vCenter has a wider range of options for managing the infrastructure, rather than low-level subsystem VMware ESXi;

2. Openstack Controller -– virtual machine based on Ubuntu Linux 14.04 LTS, which is designed for the management and coordination of environment of Openstack. Openstack – open technology, which simplifies and standardizes access to and control of virtualization;
3. Openstack Compute -– virtual machine based on Ubuntu Linux 14.04 LTS, which is designed to control a virtual cluster, in this case - the cluster under control of VMware vCenter.

The process of creating a virtual machine on the client's request includes following steps:

1. The client fills out a form on the Geoportal and determines the appropriate settings of the virtual machine (type of the operating system, the parameters of the virtual hardware) that are sent to the virtual machine Openstack Controller using the Openstack API.
2. Openstack Controller analyzes provided parameters and determines the corresponding node of Openstack Compute, which will host the created virtual machine. As currently there is only one Compute node , associated with the VMware vCenter, the task of creating a virtual machine is sent to him.
3. Openstack Compute initiates the creation of a virtual machine using the API VMware vCenter according to user-defined parameters.
4. VMware vCenter creates a virtual machine, referring directly to the ESXi subsystem. Then it returns the necessary information about new virtual machine back to the client all the way up of the virtual machine creation chain.

Note that for security reasons virtual machines are available only from the internal network, however, client virtual machines have external static IP-addresses, thus, WPS services deployed on the client virtual machines are accessible from anywhere on the Internet.

Virtual machines are created from special templates (in the terminology of VMware it is a "template in Openstack this is "image"), where Zoo Project WPS Platform and 52 ° North WPS are pre-installed and configured. ZOO Project Server allows users to publish algorithms written in different languages (Java, C / C ++, Python and Perl). 52 ° North WPS supports algorithms of GIS GRASS GIS, Sextante, ArcGIS.

Zoo Project Server was modified — the support of console applications was added. Integration of different libraries and software packages into the Zoo Project is defined by configuration files, console applications have additional run template setting. The template includes guidelines for passing parameters. On the basis of template Zoo Project forms string to startup of the console application. Console application works with the data as usual - with the files on the local file system.

Users can install theirs own software, configure access through WPS, using Zoo Project extension. In the final step, the user registers the service in the WPS Geoportal catalog:

1. entering service name, description, data to access the WPS service;
2. requesting metadata of available services, according to the entered address;
3. performing a request for metadata about the parameters of the desired service;
4. entering the additional information about the service input and output parameters: a form control which will be drawn for specific parameter, human-readable name of the parameter, explanatory text. This information is used for generation of the user interface for the parameter and the verification of entered parameter data.

After the registration the WPS service can be run using the Geoportal. It is possible to specify files from the data storage system of Geoportal as input data of the service, which will be automatically transferred to the virtual machines, and stored in the local file system. When the console application runs the transferred files are provided as local files. The file results of

the application are transferred back and stored in the storage system. The advantages of this technology are:

- the minimum knowledge and effort to create a WPS-services;
- the freedom to install the necessary software on virtual machine;
- the possibility of long-term preservation of WPS-services by storing working copies of virtual machines;
- providing computing resources permanently available on the Internet.

3 Technology for Creating a WPS Service Running on Supercomputers

With the development of parallel computing and supercomputers software systems requiring high performance are designed, implementing unique methods, techniques, data processing and analysis. Now, these methods are localized and mainly used by only developers. Therefore organizing access to such software systems as services is required, which expands the range of potential users and applications.

Technology for creating WPS services based on supercomputers uses Zoo Project WPS Platform for presenting supercomputer-based services as regular WPS services. The computer cluster "Academician VM Matrosov"is used (ISDCT SB RAS, total peak performance - 33,7 TFlops and reached a maximum performance on tests of HPL - 25,12 TFlops) as the computing platform. This cluster provides a REST-interface that allows user to queue, run and track tasks.

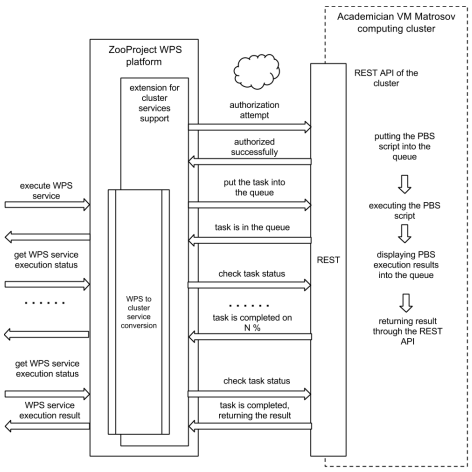

Fig. 2. Scheme of access to the cluster software systems using WPS standard.

In order to perform the WPS service on the computing cluster the service should be pushed to the cluster queue. The special extension of Zoo Project has been developed, which is based on the configuration file that is used for executing queries to the REST interface of the cluster. Cluster-based service execution and processing of its results is carried out in following steps:

1. The user enters initial parameters of the WPS-service in one of the WPS client and starts.
2. Zoo Project initiates the execution of a long process, and makes a request to the REST-interface of the computing cluster.
3. The task is queued for execution on the cluster, Zoo Project periodically checks the status of the task and informs the client (user).

4. After completion of the task Zoo Project returns results in accordance with WPS.

Thus, the main goal is achieved -— integration of computing environment of the cluster software systems of "Academician VM Matrosov"as the set of WPS services.

3.1 Program Code of Service Combination

As an example, the road_pnt_pol scenario will be described. This scenario calculates pollution of point sources and roads in cells of the regular grid. The analyzed scenario uses several distributed WPS-services – vector2grid, road2grid, g_sum running on virtual machine and the supercomputer.

Example of Scenario

```
function road_pnt_pol(housefile, roadsource, commonresult, sumpol){
       var houseresult='/tmp/hr.tif';
       var roadResult ='/tmp/hr.tif';
       vector2grid(housefile, houseresult);
       road2grid(roadsource, roadResult, sumpol);
       g\_sum(roadResult, houseresult, commonresult);
       return true;
}
```

(Example of Scenario written in JavaScript.)

WPS-service vector2grid calculates pollution from point sources. Vector points file, that describes the location and amount of emitted pollution, serves as an input for this service. GeoTIFF file is created as a result of vector2grid execution, it displays summarized pollution in every cell of the grid. WPS-service road2grid calculates pollution from the roads. Vector file, that describes location of roads, theirs average load and summarized amount of pollution for the settlement. Result of road2grid is the GeoTIFF file that describes summarized pollution in every cell. WPS-service g_sum takes two GeoTIFF's that were generated as a result of execution before and combines them in the new file.

In this example, components of the environment of distributed WPS-services ensure following:

1. scenario execution;
2. passing remote WPS-services file data from the storage system, defined as input parameters;
3. receiving result data and placing it in the storage system using predefined pathes;
4. showing executing state, displaying WPS services messages.

Automation of WPS service invocation using JavaScript reduces time of the calculation for the variety of air quality improvements, minimizes number of user errors.

4 Conclusion

The proposed technologies allow to simplify service creation in cloud environment, supercomputers and to combine distributed program systems into single application. The advantages of these technologies are:

– interaction of program systems running on various platforms;
– sending and receiving data among services;
– simplifying development of program systems in accordance with WPS standard.

References

1. Papazoglou, Mike P.; van den Heuvel, Willem-Jan, "Service oriented architectures: approaches, technologies and research issues", VLDB JOURNAL, Volume: 16 Issue: 3 Pages: 389-415 Published: JUL 2007
2. *OpenGIS Web Processing Service (WPS) Implementation Specification*, v1.0.0. Release date: June 08, 2007. – URL: http://www.opengeospatial.org/standards/wps (04.09.2014).
3. X. Yang and H. Zhang, "Cloud computing and SOA convergence research," in Proceedings of the 5th International Symposium on Computational Intelligence and Design (ISCID '12), vol. 1, pp. 330–335, Hangzhou, China, 2012.
4. Bychkov I., Plyusnin V. M., Ruzhnikov G., Fedorov R., Hmelnov A., Gachenko A. *The creation of a spatial data infrastructure in Management of regions (exemplified by Irkutsk oblast)*, Geography and Natural Resources. - 2013. - Vol. 34. No2. - P. 191 - 195.

Synthesis of Plans in Multi-Agent System Using the Method of Positively Constructed Formulas

Artem Davydov*[1], Alexander Larionov[2], and Ivan Terekhin[3]

[1] Matrosov Institute for System Dynamics and Control Theory of SB RAS, Irkutsk, Russia, artem@icc.ru
[2] Irkutsk State University, Irkutsk, Russia, alexander@prisnif.org
[3] Irkutsk State University, Irkutsk, Russia, i.terhin@gmail.com

Abstract. In this paper, we consider the use of the logical calculus of positively-constructed formulas (PCF) for making plans of actions for groups of autonomous underwater vehicles (AUVs). Actions of groups and group members are generated at the high level of control system in automatic or semi-automatic mode via the human-machine interface. We presenting an example of formalization of such system and the construction of inferences.

Keywords: automated theorem proving, intelligent control, multi-agent systems.

1 Introduction

The calculus of positively constructed formulas (PCF) was presented in [6,2] by S.N. Vassilyev and A.K. Zherlov. The PCF is a first-order logical formalism. The PCF calculus originally developed during the solving of control theory problems. Area of this problems and the approaches also presented in [2]. Usability of the calculus due to some features (that will be described below) allow ones to combine automatic proof with special problem heuristics (human knowledge and experience), therefore PCF calculus can be positioned both as machine-oriented and human-oriented. PCF calculus from automated theorem proving (ATP) point of view, its soundness and completeness, proof search strategies and their implementation are described in detail in [2]. Here, we present the extended logical calculus of positively-constructed formulas, with functional symbols. In this work, in contrast to [6,2] PCF calculus presented in a slightly different notation more close to a modern, used in the literature on related topics.

The approach to the implementation of the high-level control of a groups of autonomous underwater vehicles based on the formalization of the domain via PCF calculus is considered. The logical deductions of PCF fragments formalizing the process of functioning of the groups of AUVs controlled by some supervised control system showed as the example.

2 PCF Calculus

At first, let us consider the language and the calculus of PCF.

2.1 The Language of PCF

As mentioned above, we consider the first-order logic language. Formulas of this language are built from atoms, logical operations $\&, \vee, \neg, \rightarrow, \leftrightarrow$, quantifier symbols $\forall$ and $\exists$, and constant operators $true$ and $false$. The concepts of variable, term, atom are understood as in the classical first order logic (FOL).

* The reported study was partly funded by RFBR according to the research projects No. 14-07-00740-a, 14-01-31449-mol-a.

Let $X = \{x_1, \ldots, x_k\}$ is the set of variables, $A = \{A_1, \ldots, A_m\}$ is the set of atoms, and $F = \{F_1, \ldots, F_n\}$ is the set of some formulas of FOL. Then, the following formulas

$$((\forall x_1) \ldots (\forall x_k)(A_1 \& \ldots \& A_m \to (F_1 \vee \ldots \vee F_n)))$$

$$((\exists x_1) \ldots (\exists x_k)(A_1 \& \ldots \& A_m \& (F_1 \& \ldots \& F_n)))$$

will be written as $\forall_X A\colon F$ and $\exists_X A\colon F$ respectively, the quantifier $\forall$ corresponds to $\to F$, where F denotes the disjunction of all formulas in F, and quantifier $\exists$ corresponds to $\&F$, where F denotes the conjunction of all of the formulas in F.

If $F = \emptyset$, then the above formulas will have the form $\forall_X A\colon \emptyset \equiv \forall_X A \to false$ and $\exists_X A\colon \emptyset \equiv \exists_X A \& true$, as the empty disjunction of elements is identical to $false$, while the empty conjunction of elements is identical to $true$. Notations for $\forall_X A$ and $\exists_X A$ are abbreviations of the above formulas. If $X = \emptyset$, then $\forall A\colon F$ and $\exists A\colon F$ are similar abbreviations.

The set A is called *conjunct*. Empty conjunct is identical to $true$, as mentioned above.

The variables from the set X bound with corresponding quantifiers are called $\forall$-variables and $\exists$-variables respectively. In the notation $\forall_X A$ the variable from the set X that is not a member of the conjunct A is called *unconfined* variable. The following abbreviations will also be used: $\forall\emptyset \equiv \forall\emptyset\colon \emptyset \equiv \forall true \to false \equiv false$.

The constructs $\forall_X A$ and $\exists_X A$ are called positive type quantifiers because A is a conjunction of positive atoms understood as type condition for X. In fact, these structures denote the following phrases: " for every X satisfying A holds ... " or " there is X satisfying the property of A such that ...". For example, "for any integers x, y, z and $n > 2$ there is $x^n + y^n \; nez^n$".

Originally, the term "type" quantifier was proposed by Bourbaki [4] as part of the notation for the formalization of mathematics. However, type quantifiers generally accepted in other related fields.

Definition 1 (Positively constructed formulas (PCF)). *Let, X is a set of variables, A - conjunct.*

1. *$\exists_X A$ and $\forall_X A$ are $\exists$-PCF and $\forall$-PCF respectively.*
2. *If $F = \{F_1, \ldots, F_n\}$ are $\forall$-PCF, then $\exists_X A\colon F$ is an $\exists$-PCF.*
3. *If $F = \{F_1, \; ldots, F_n\}$ are $\exists$-PCF, then $\forall_X A\colon F$ is a $\forall$-PCF.*
4. *Any of the $\exists$-PCF or $\forall$-PCF is PCF.*

This type of logical formulas called positive-constructed formulas, as they are recorded only through positive type quantifiers. The formulas does not explicitly contain the sign of negation. Any formula of FOL language can be represented as a PCF [2]. Thus a PCF is a special form of notation for the classical formulas of FOL like the conjunctive normal form, disjunctive normal form and others.

PCF starting with $\forall\emptyset$ called the PCF in canonical form. Any PCF can be represented in a canonical form. Indeed, let F is a $\exists$-PCF, then the canonical formula $\forall\emptyset\colon F$ is the equivalent representation of F, because: $\forall\emptyset\colon F \equiv true \to F \equiv F$. If by opposition F is not a canonical $\forall$-PCF, then $\forall\emptyset\colon \{\exists\emptyset\colon F\}$ is its canonical representation since $\forall\emptyset\colon \{\exists\emptyset\colon F\} \equiv true \to \{true \& F\} \equiv F$. Type quantifiers $\forall\emptyset$ and $\exists\emptyset$ are called *fictive*, since they do not affect the truth value of the initial PCF and do not bind any variables, but are serving only as a structures maintaining a correct notation of PCF.

For the ease of formula's readability we will represent them in the form of a tree as follows:

$$Q_X A\colon \{F_1, \ldots, F_n\} \equiv Q_X A \begin{cases} F_1 \\ \ldots \\ F_n \end{cases},$$

where Q – some quantifier. Tree elements are called as usual: node, root, leaves, branches, etc. Since the quantifiers $\forall$ are corresponding to disjunction in the formulas $\{F_1, \ldots, F_n\}$, (quantifiers $\exists$ are corresponding to the conjunction), then, in general, all $\forall$-nodes are corresponding to *disjunctive branching*, and $\exists$-nodes are corresponding to *conjunctive branching*.

Some parts of canonical PCF are called as follows:

1. root node of PCF $\forall\emptyset$ will be called *PCF root*;
2. nodes $exists_X A$ that are the children to the root of PCF called *bases of PCF*, conjunct A is called em base of the facts, and any PCF wich root node is the base called em base subformula;
3. nodes $\forall_Y B$ that are the children to the base of PCF called em questions to the appropriate base. If the question is a leaf of the tree, it is called *the goal question*.
4. Subtrees of the questions are called em consequents. Consequent of the goal question is $false$.

Example 1. Let us consider a formula of classic first order logic:

$$F = \neg\big(\forall x\, \exists y P(x,y) \to \exists z P(z,z)\big).$$

An image F^{PCF} of F in the PCF language is

$$F^{\text{PCF}} = \forall\colon \boldsymbol{True} - \exists\colon \boldsymbol{True} \begin{cases} \forall x\colon \boldsymbol{True} \ - \exists y\colon P(x,y) \\ \forall z\colon P(z,z) - \exists\colon \boldsymbol{False}. \end{cases}$$

We will consider PCFs, in general, in the canonical form. The image of any canonical PCF in the FOL language is equivalent to the disjunction of images of its base subformulas. Since the method of automatic deduction of PCF is working by refutation of statement negation to be proved, the refutation of the disjunction is reducing to independent refutations of images of each base subformulas. For that reason, Herbrand universe as a universal interpretations domain will be defined only for base subformulas. Also, it should be noted, that variables bound by a base quantifier in the formula act as constants since instead of $\exists$-variables one can not substitute any terms. Moreover, without loss of generality, we can consider these variables as constants, as by the well-known theorem about Skolemization, the eliminating of the existential quantifier and replacing it with a new constant, that is not occurred in the formula, does not affect the inconsistency of the original formula.

Definition 2. *Let $G_1, \ldots, G_k$ — basic subformulas of a PCF F, $G_i = \exists \bar{x}_i\colon B_i(\tilde{x}_i)(\widetilde{\Phi})$, $i = \overline{1,k}$. Let associate for each G_i the set $H_i^0, H_i^1, \ldots, H_i^\infty$ as follows: $H_i^0 = \{\ barx_i \cup \bar{a}_i\}$, where $\bar{a}_i$ — a list of all constants occurring in G_i; if H_i^n has been built $(n \geqslant 0)$, then H_i^{n+1} is the union of H_i^n with a set of all kinds of terms of the form $\bar{f}(t_1, \ldots, t_r)$, for all functional symbols $\bar{f}$, occurring in G_i, and all lists $(t_1,\ ldots, t_r)$ of terms of length r equals to the arity of $\bar{f}$ (thus, these terms are constructed in the basis of elements from H_i^0 and functional symbols of G_i). Every set H_i^j called Herbrand subset of the j-th level, corresponding to G_i. The set H_i^∞ is defined by $H_i^\infty = \bigcup\limits_{j=0}^{\infty} H_i^j$ and is called Herbrand set corresponding to G_i (Herbrand universe of G_i).*

Example 2. In the following formula:

$$\forall\emptyset \ - \ \exists\colon Man(Socrates) \begin{cases} \forall x\colon Man(x) & - \exists\colon Mortal(x) \\ \forall\colon Mortal(Socrates) & - \exists\colon false \end{cases}$$

there is one constant - *Socrates*. To the only base subformula of the given formula is corresponding the Herbrand set: $H^\infty = \{Socrates\}$.

Example 3. In the formula:

$$\forall\emptyset \ - \ \exists x, y, z\colon P(x, y, z) \ - \ \forall x_1\colon P(e, f(e), e) \ - \ \exists\colon Q(x_1)$$

there is one constant e, one function symbol f. To the only base subformula of the given formula is corresponding the Herbrand set: $H^{\infty} = \{e, f(e), f(f(e)), \ldots x, f(x), f(f(x)), \ldots y, f(y), f(f(y)), \ldots z, f(z), f(f(z)), \ldots\}$.

Definition 3. *Substitution theta is the following set: $\theta = \{x_1 \to t_1, \ldots, x_n \to t_n\}$, where x_i - variable, t_i - terms, and in all t_i not occurring $x_i, i =$ overline$1, n$, foralli neqj colonx_i neqx_j. It is said that θ defined on the set of variables $\{x_1,$ ldots$, x_n\}$. Substitutions will be denoted as well as a mapping: $\theta\colon \bar{x} \to T$, where x is set on which is defined this substitution, T is a set of terms.*

2.2 Inference rule

Definition 4 (Answer). *It is said that some question $\forall_Y B : \Upsilon$ to the base $\exists_X A$ has an* answer *θ if and only if θ is a substitution $Y \to H^{\infty} \cup X$ and $B\theta \subseteq A$, where H^{∞} is the Herbrand universe corresponding to the given base subformula.*

Definition 5 (Split operation). *Let $F = \exists_X A\colon \Psi$, and $S = \forall_Y B\colon \Upsilon$, where $\Upsilon = \{\exists_{Z_1} C_1\colon \Gamma_1, \ldots, \exists_{Z_n} C_n\colon \Gamma_n\}$ then $split(F, S) = \{\exists_{X \cup Z_1'} A \cup C_1'\colon \Psi \cup \Gamma_1', \ldots, \exists_{X \cup Z_n'} A \cup C_n'\colon \Psi \cup \Gamma_n'\}$, where $'$ is the variables rename operator. It is said that F is splitting by a question S. Obviously, $split(F, \forall_Y B) = split(F, \forall_Y B\colon \emptyset) = \emptyset$.*

Definition 6 (Inference rule ω). *Let us consider a canonical PCF $F = \forall\emptyset\colon \Phi$. Suppose that there is a question Q that has the answer θ to a base $S \in \Phi$, then $\omega F = \forall\emptyset\colon \Phi \setminus \{S\} \cup split(S, Q\theta)$.*

In the other words, if the question has the answer to its base, then the base subformula is splitting by this question. In the case when the goal question has an answer, it is said that the base subformula refuted because $split(S, \forall_Y B) = \emptyset$. Refuted base subformula S is removing from the set of base subformulas Φ, since $\Phi \setminus \{S\} \cup \emptyset = \Phi \setminus \{S\}$.

As soon as all the base subformulas of Φ are refuted, F formula will be refuted too, because $\forall\emptyset\colon \emptyset \equiv false$ according to the rule described above.

Note that the answer to the question with the disjunctive branching increases the number of base subformulas due to the use of the split operator, and refutation of a base is reducing the number of bases by one just refuted.

Example 4. A refutation in $\boldsymbol{JF}$

$$F_1 = \forall\colon \boldsymbol{True} - \exists\colon S(e)(Q_1, Q_2, Q_3, Q_4);$$

$$Q_1 = \forall x\colon S(x) - \exists\colon A(a);$$
$$Q_2 = \forall x, y\colon C(x), D(y) - \exists\colon \boldsymbol{False};$$
$$Q_3 = \forall x, y\colon B(x), C(f(y)) - \exists\colon \boldsymbol{False};$$
$$Q_4 = \forall x\colon A(x) - \begin{cases} \exists y\colon B(y), C(f(x)) \\ \exists\colon C(x) - \forall z\colon A(z), C(z) - \exists\colon D(f(z)). \end{cases}$$

At the first step there exists only one answer $\{x \to e\}$ to question Q_1. After applying rule ω with this answer, formula will have the following form:

$$F_2 = \forall\colon \boldsymbol{True} - \exists\colon S(e), A(a)(Q_1, Q_2, Q_3, Q_4).$$

At the second step there exists only one answer $\{x \to a\}$ to question Q_4. After applying ω with this answer, formula is split, because Q_4 has disjunctive branching. And now formula has the following form:

$$F_3 = \forall : \boldsymbol{True} - \begin{cases} \exists y_1 \colon S(e), A(a), B(y_1), C(f(a)) - \begin{cases} Q_1 \\ \cdots \\ Q_4 \end{cases} \\ \exists \colon S(e)A(a), C(a) - \begin{cases} Q_1 \\ \cdots \\ Q_4 \\ \forall z \colon A(z), C(z) - \exists \colon D(f(z)). \end{cases} \end{cases}$$

At the third step first base can be refuted by answering on Q_3 goal question with $\{x \to y_1; y \to a\}$. Refuted base (base subformula) can be deleted from the list of base subformulas.

At the fourth step there exists the answer $\{z \to a\}$ to fifth new question. And now formula has the following form:

$$F_4 = \forall \colon \boldsymbol{True} - \exists \colon S(e), A(a), C(a), D(f(a)) \begin{cases} Q_1 \\ \cdots \\ Q_4 \\ \forall z \colon A(z), C(z) - \exists \colon D(f(z)). \end{cases}$$

At the fifth step the only base can be refuted by answering on Q_4 goal question with $\{x \to a; y \to f(a)\}$.

The refutation is finished because all bases were refuted.

2.3 Features of PCF Calculus

1. Any PCF has a *large-block structure* and only *positive quantifiers* $\exists$ and $\forall$.
2. PCF contains both $\exists$ and $\forall$ quantifiers, but structure of PCF is *simple*, *regular* and *predictable* due to predictable quantifiers alternation $\exists$ and $\forall$ in all the branches.
3. There is no need to preprocess (Skolemization) the original formulas by elimination of all existential quantifiers. The Skolemization related to this elimination leads to elevating the complexity of terms. Also, this feature makes PCF calculus more human-oriented.
4. "Theoretical"quantifiers $\forall x$ and $\exists x$ usually not used in the formalization of human knowledge, typical quantifiers $\forall x(A \to \sqcup)$ and $\exists x(A\&\sqcup)$ are used instead.
5. The PCF calculus inference rule is unique and this property make PCF calculus more machine-oriented. In addition, inference rule has a large-block structure that make PCF calculus more human-oriented.
6. The proof procedure focuses on the root surroundings of the PCF. This is possible due to features 1, 2.
7. The proof procedure can be represented in terms of the *question-answering* procedure instead of technical terms of formal deducibility (i.e., in terms of logic connectives, atoms, etc.). Base conjunct can be interpreted as *base of facts*.
8. There is natural *OR-parallelism*, because the refutation of base subformulas are executed independently of one another.
9. Due to 1, 2, 6, 7 the proof procedure is *well compatible with heuristics*, as well as to a general proof control heuristics. Due to feature 5 proof consist of large-block steps, and it is well *observable* and *controllable*.

10. Due to features 7, 9 the proof obtained are quite interpretable by human. This interpretability of proof is quite important from the viewpoint of human-machine applications. Thus, as mentioned above, PCF calculus are not only machine-oriented, but also human-oriented.
11. Semantic of PCF calculus can be changed without any modifications of axioms and inference rule ω. Such modifications are implemented merely by some restrictions to application ω and allow us to transform classical semantics of PCF calculus in non-monotonic, intuitionistic, etc. Examples of using such semantics are provided in [2].

3 Planning and control of AUVs groups

In the beginning let us consider, without going into details, some history moments of the development of methods for solving the problems of planning based on the logical means of representation and processing of knowledge. As the methods of automated reasoning based on classical first order logic was advancing, even before the most well-known method for automated theorem proving - resolution method, developed by J. Robinson in 65, McCarthy proposed a formalism based on first-order predicate logic. Then, along with the development of temporal logics, which are increasingly used for solving problems of verification, the project STRIPS for mobile robot control was being developed in Stanford, an approach based on non-monotonic first order logic, which has been advanced to an independent, more general formalism - Action Definition Language. In '98 was developed PDDL language as a universal means for describing planning problems. In this language, you can denote problems formalized in the planner programs, based on different formalisms and to evaluate the effectiveness of these planners. Though all this formalisms was in some way general and fitting to solve problems of planning they all have some restrictions, mainly in expressiveness of its base languages.

Due to the features of the PCF, the restrictions of the logical approach to planning and control are significantly reduced. Question-answering procedure of inference search is making possible the use of PCF to solve problems. Indeed, in this way of searching of deductions, it is easy to pick out the events that can be used to tuning up a strategies of deduction search in automatic mode. And also, with the help of these events the inference machine can communicate with a human operator of the control system. The most natural way to implement such a connection is the use of the following events:

- a successful answer to the question;
- updating the database with new atoms;
- refutation of a base.

In contrast to [2] in this work we consider the full calculus with functional symbols, that is greatly facilitate the formalization of planning and control problems which demand the declarative specification. Generally speaking, the use of function symbols leads to a complexity of the search of inferences due to the possibility of emerging the Herbrand universe while searching the substitutions needed for inference. Therefore, in order to avoid this inefficiency, we limit the language of calculus of PCFs as follows. While formalizing problems it is disallowed the use of formulas containing questions that are simultaneously contain unconfined variables and branching in the tree structure of the formula. This restriction significantly improves the efficiency of the search of inferences, but the class of formulas remains fairly extensive, transcending the class of formulas used for example in the "pure" Prolog.

Let us consider the example of calculus of positively constructed formulas application to describe and solve the following problem. Imagine a fragment of a centralized control system of a groups of autonomous underwater vehicles that are continuously monitoring certain underwater

area. In the area there are two groups of autonomous robots (a_1 and a_2). Groups are composed of robots with different functionalities united in the likeness of tasks (for example, identification, sampling and manipulation, etc.). Control of the groups is maintained by means of acoustic communication network with some central server (CS), that is located on support ship or on shore. CS operates automatically under the supervision of a human operator which is able to make changes to the now performing tasks and change the goals. Next, consider PCF language specifications of robots and CS as a possible parts of the larger system specification, focusing here on achieving a common goal of the two groups that will be reached in stages, to complete each stage it is required to achieve sub-goals. Herewith, completing or inability to complete sub-goals is defined by events in the deductions search of formulas that specify the functionality of robot groups. If there is no way to complete a sub-goal, then there is taking place a communication with a central server to request re-planning.

Sub-goals that are possible in our formulation of the problem:

- object identification;
- carrying out additional actions with the object (sampling, communication);
- conducting defensive operations;
- loading of an object;
- loading on the object;

Groups of robots specifications in some initial moment of abstract time can be described by the following base subformula:

$$\exists s_1, s_2, a_1, a_2 \colon S(s_1), A(a_1), S(s_2), A(a_2), T(0)$$

Here $S(s_i)$ denote the set of predicates describing the state of the groups $A(a_i)$.

Current functioning and tasks for groups describe the following questions to the base subformula:

$$\begin{array}{ll} \forall t : T^*(t) & - \exists \colon T(t+1) \\ \forall x, t \colon Patrol(t, s_i, a_i), T(t), See^{\#}(x) & - \exists \colon Find(t, x) \\ \forall x, t \colon Find(t, x) & - \exists \colon \boxed{Task(t, x, a_i)} \end{array}$$

First in the list of questions is our method of time counting in PCFs. It uses the atom with the symbol $*$, which means that if there is a successful answer to a question, one must delete the corresponding atom from the base of facts. Thus from the initial moment of time in cycle of question bypassing in the base is added T(1) instead of T(0). Further, if the base is not refuted, T(2) and so on.

Second question. Assume that one of the groups has the capability to detect objects. This option is checked in cycle of question bypassing with the help of computational predicate $See(x)$, denoted hereinafter in formulas by the symbol #. The truth values of the computational atoms are not establishing by a logical inference, but by the actual data of environment. Thus, if detection sensors trigger successfully, the fact $See(Obj)$ - a symbolized definition of the detected object - is added to the base. Then, for example, at time moment n, the second question has an answer, and the fact $Find(n, Obj)$ is adding to the base.

Framed atom in the third question indicates the achievement of one of the sub-goals and the end of the deduction search, further there is a communication with the CS. So in our case, by finding the object there is a request to the CS for further instructions, the atom $Task(n, Obj, a_i)$ is sending to the base of facts of the CS.

For this behavior, in the list of questions on the CS might present the following question:

$$\forall t, x \colon Task(t, x, a_i) - \exists \colon \boxed{Identify(x, a_i)}$$

When answering to this question, the fact $Identify(Obj, a_i)$ is sending to the coordinator of the group sent the request.

Further work of the system is related with the study of the object detected. The first step is to determine whether the found object can present any danger. Suppose that by the functionality a given task can handle group a_1, if there is no danger, the group a_2 can proceed to the task, which have robots with the ability to capture photos and videos.

Any object found in the area can be considered hazardous if it is moving and is not a local fauna, or if it contains any hazardous materials (pollution, underwater mines). In the list of questions of group coordinator of a_1 for this task there must be the following:

$$\forall x\colon Identify(x, a_1) \begin{cases} \exists\colon Moving^{\#}(x), NoAnswer^{\#}(x), NotFish^{\#}(x) \\ \exists\colon Contamination^{\#}(x) \\ \exists\colon Mine^{\#}(x) \end{cases}$$

$$\forall x\colon Moving(x), NoAnswer(x), NotFish(x) - \exists\colon \boxed{Danger_1(x, a_1)}$$

$$\forall x\colon Contamination(x) - \exists\colon \boxed{Danger_2(x, a_1)}$$

$$\forall x\colon Mine(x) - \exists\colon \boxed{Danger_3(x, a_1)}$$

The first question in this list is splitting the base, each of which will correspond to the type of danger. Atoms with the symbol # request the corresponding sensors. In case of a negative analysis of the object corresponding atoms in the base are replaced by a constant *true*. According to the problem statement, the answer can exist only to one of the last three questions from the list, if it exists, deduction search stops and type of danger - $Danger_j(Obj, a_1)$ - is sent to the CS. If the deduction search fails, the group a_1 returns to normal functioning, and the CS sends to group coordinator of a_2 task for the photo and video shooting, which processes it with the help of questions:

$$\forall x, t\colon PhotoVideo(x, a_1), T(t) \begin{cases} \exists\colon Photo(x, t) \\ \exists\colon Video(x, t+1) \end{cases}$$

$$\forall x, t\colon Photo^{\#}(x, t), T(t) - \exists\colon false$$

$$\forall x, t\colon Video(x, t+1), T(t) - \exists\colon false$$

The deduction search of the two new bases will be successful, one after another. First, at some moment T(n) there will be a photo shoot with the help of computational atom of one from the bottom question in the list, then, after bypassing of one cycle in the second base time moment will change to T (n + 1) and it will be possible to answer the last question, refuting base and thereby completing the deduction search. End of the deduction search means the achieving of another sub-goal.

Further, by analogy, it is possible to describe all possible situations and circumstances of the environment of the area in the system. As described above is carried out “life cycle” of two groups of robots. During the deductions search of formulas describing multi-agent system, in automatic or semi-automatic mode plans of action are building and performing. This example demonstrates a common approach to formalizing and solving the planning problem using the method PCF.

4 Conclusion

In the report the extended logical calculus of positively-constructed formulas, with functional symbols, its features and examples of deductions search is considered. In this work, in contrast to [2,6] PCF calculus presented in a slightly different notation more close to a modern, used in the literature on related topics.

The approach to the implementation of the high-level control of a groups of autonomous underwater vehicles based on the formalization of the domain via PCF calculus is considered. The logical deductions of PCF fragments formalizing the process of functioning of the groups of AUVs controlled by some supervised control system showed as the example.

References

1. Vassilyev, S.N.: Machine Synthesis of Mathematical Theorems. The Journal of Logic programming, V.9, No.2–3, pp. 235–266 (1990)
2. Vassilyev, S.N., Zherlov, A.K., Fedunov, E.A., Fedosov, B.E.: Intelligent Control of Dynamic Systems. Moscow: Fizmatlit (2000)(in Russian)
3. Davydov, A.V., Larionov, A.A., Cherkashin, E.A.: On the calculus of positively constructed formulas for automated theorem proving. Automatic Control and Computer Sciences (AC&CS), V.45, No.7, pp. 402–407 (2011)
4. Bourbaki, N.: Theory of Sets. Paris, Hermann (1968)

Algorithm for Recognition of Kazakhstan Vehicle License Plates

Beibut Amirgaliyev, Magzhan Kairanbay, Chingiz Kenshimov, Kanat Kuatov, Zhalgas Baibatyr, and Abylay Jantassov

ABY Applied System,
Abay avenue. 52V, 050008 Almaty, Kazakhstan
amirgaliyev@gmail.com,magzhan.kairanbay@gmail.com,kenshimovchingiz@gmail.com,kkuatov@yandex.ru,
presscorp@yandex.ru,abylaikz@gmail.com
http://www.abyas.kz

Abstract. This paper presents Automatic license plate extraction, character segmentation and recognition for Kazakh vehicles. Recently Kazakhstan provided new format of the number plate in addition to old format numbers, which differ from previous one. Due to variations in the representation of number plates, vehicle number plate recognition is crucial. We provide own algorithm to recognize new and old formats of Kazakhstan number plates. A short review is performed on the various methods for automatic license plate recognition (ALPR) algorithms. Further explanations of the proposed algorithm are illustrated in graphical forms to show how the algorithm works. Camera localization and new approach of segmentation is explained in detail taking into account of parameters. This article concluded with tests and evaluation results.

Keywords: license plate localization, optical character recognition, segmentation.

1 Introduction

In the new global economy, traffic congestion became a central issue for majority of the developing countries. The number of cars are increasing rapidly; respectively the numbers of violations are increasing. Speeding, stealing the cars and other wide spectrum of violations on the road are general things in our daily lives. Shortages of parking places are reasons to entering unauthorized cars to the private areas and spending a lot of time to find free places in parking lots.

Automated number plate recognition system is a key aspect in resolving all the problems listed before.

Kazakhstan introduced its own standard format of number plate in 1993.However, at the end of 2012, Kazakhstan begun to provide new format of number plates, which is significantly different from previous one. The old formats of number plate are rectangular or square shaped with black colored characters on the white board (Fig. 1). First character identifies the place of registration. Total number of characters depends on registration to legal entity or an individual. For vehicles with a non-adjustable mounting location and for the most of cars made in Japan, have been prescribed variants with square license plates. The second generation of vehicle license plates remains with the same principle, in terms of total number of characters (Fig. 2). However, added Kazakhstan flag and index KZ on the left part of the plate. The style of font changed, too.

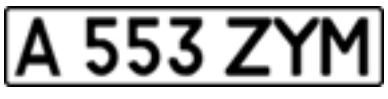

Fig. 1. Old format of Kazakh vehicle numbers

Fig. 2. New format of Kazakh vehicle numbers

In this paper we are going to review such algorithms which will be able to recognize the both new and old formats of Kazakhstan number plates.

1.1 Related Work

There are various solutions of relevant problems. The main issues in number plate recognition are climate conditions, environmental interference, and accuracy of number plate localization. One of the methods of recognizing the number plate is utilizing the color characteristics and probability distribution of the license plate between the two lights [1]. Another popular method of number plate recognition algorithm is template matching [2]. The License Plate Detection algorithm based on template matching was designed and written for managing the parking lot system by identifying the unregistered cars from off- campus. At the same time vertical edges-based car license plate detection [3] are popular, too. However, others prefer to find the location number plate using horizontal and vertical projections of image [4]. The Genetic Algorithm [5] and Hough transform [6] can be applied to detect the license plate area. At the same time, the combination of edge statistics and mathematical morphology showed good results [7], [8] and they use block-based algorithm. Another algorithm [9], which is based on rows distances counts the existent edges and if this number is more than some threshold value then number plate is recognized. Wavelet transform-based algorithm extract the important features to be used for number plate localization. The advantages of this algorithm, it will allow to find more than one number plates in the frame. Some of the methods above are very complex and requires too much computation time, which is a bit difficult to use in real time applications. However, other approaches could be used only in specific countries with specific characteristics of number plate like background color, etc.

1.2 Research Objectives

The following list gives the objectives of this research paper:

1. Solve and develop the automated number plate recognition system for Kazakhstan vehicle license plates
2. Calculate the localization parameters of camera taking into account the format of Kazakhstan vehicle license plates
3. To evaluate, test developed system and presents the evaluation results

2 Proposed Solution

The overall problem consists of three parts:

1. Plate area detection
2. Segmentation or extraction of characters from number plate
3. Optical character recognition (OCR) of extracted symbols

2.1 Plate Area Detection

The first part of algorithm should find the location of number plate. The image with number plate will be given as an input to the program and the number plate must be identified then cropped as output image to the next stage. Detection of number plate is one of the critical stages in the overall system, because it directly effects to the segmentation and optical character recognition stages. In order to determine the number plate from whole image, firstly we will convert the image to gray-scale format then blur it. The characters inside the number plate contain mostly vertical edges in comparison to horizontal. Therefore, one of the best approaches is to find vertical edges that are too close to each other [10]. The edge detection is basic and fundamental operations in computer vision. There are different kind of edge detectors like Prewitt, Sobel, Canny and etc. Each of them is used in different cases and problems. We use Prewitt, Sobel and modified version of Sobel which based on gradient magnitude and direction [11]. However, after investigation and testing we came to conclusion to use the modified version of Sobel, because it correctly identifies vertical edges and reduce the most of horizontal edges that impede to find the number plate. After finding vertical edges we will apply Otsu threshold to our image. Since our goal is to find the area of number plate, we do not need much information about characters. So we apply close morphology, where all letters and digits combined together. After applying the close morphology, we will find those contours that look like number plate, where the area and aspect ratio of contour must be taken into account. In order to find the correct candidate among others, first, we should compute the feature vector for each candidate. In order to do that we use Histogram of oriented Gradients, which creates feature vector for each candidate based on its gradients. Giving all of these vectors to the Support Vector Machine (SVM) and setting it with specific settings we will be able to obtain the correct number plate among others.

2.2 Segmentation

Characters must be extracted from number plate. There are two basic algorithms for segmentation, where first one is based on projection of image into X axis, however, the second one based on finding of contours that look like character. The main issue in segmentation stage is binarization. Due to the presence of sun light, shadows and other conditions, the characters inside the license plate mostly combined together which will lead to unsatisfactory of character extraction from number plate. For that case, we introduce simple method for elimination of upper and lower shadows. It was done by increasing the binarization threshold by some dynamic value. During the tests we took this value equal to 0.6*sigma where sigma is the standard deviation. In the research [12] which made for recognition of ancient documents, Nick algorithm was used for binarization. During the testing our algorithm it turned out that Niblack binarization algorithm with different coefficient equal to -0.65 and window size (both height and width) equal to the height of number plate works better than original algorithm. According to investigation and testing, we came to conclusion to use second algorithm which works better in comparison to first one (Fig. 3).

Fig. 3. The extraction of character using contour based method

2.3 Optical Character Recognition

Extracted character must be recognized. For recognition we used KNN and modified version of 1NN algorithm where character was divided into small 49 sub-parts (Fig. 4).

Fig. 4. Division of image into small 49 sub-parts

Each subpart of images should be numbered according to occurrence of white colored pixels. The feature vector with 49 features will identify each character (Fig. 5). The image should be

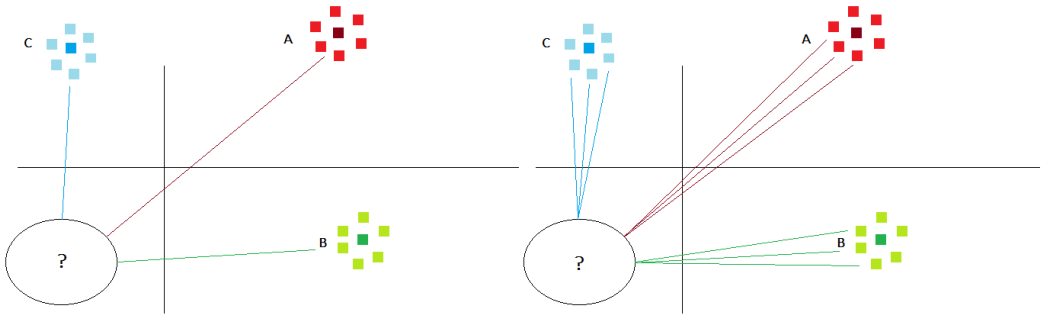

Fig. 5. Modified version of 1NN and KNN algorithms

numbered for each average element of class based on feature vectors. For unknown element the distance to all average elements of each class must be calculated. Unknown element will be joined to those neighbor class that is closest to that element. KNN works better and longer in comparison to modified version of 1NN. Instead of calculating the distance to average elements of each class, we will calculate all distances and sort it. Unknown element will be joined to those neighbor class that is appears frequently in first K sorted distances. After testing, we came to the conclusion that the KNN works better in comparison to modified version of 1NN.

3 Camera Localization

The camera that we used for capturing the frames is CCD Camera, model: P6-HM6306R (1.3MP 720P HD IP IR camera). According to the CCTV information from the CCTV Advisory Service [13], the plates of UK have a retro reflective property, where reflected light will come back to light source. The Kazakhstan vehicle license plates have such property, too. Irrespective to wide spectrum of lighting conditions like daytime, night-time, sunlight, there should be constant level and direction of illumination. Changing the contrast of camera and providing natural and artificial light source we will be able to remove noisy elements and get only plate in the frame.

Kazakhstan vehicle license plates are 520mm long x 112mm high. The height of characters approximately 77mm. According to the research [13] which made for UK license plates, the

best position for illuminator and single camera covering barrier entrance in a 1M high bollard viewing directly at the approaching vehicle. Based on same research the plate should be 18% of the scene width or vertical height of the character should be 3% of 400 line camera. Taking into account this fact, the scene height will be equal to 2566mm, the distance from camera to vehicle approximately 3500mm and lens angle 20.65 (Fig.6). Setting up the shutter speed to 1 / 1000th second we will be able to increase OCR translation in blurred images.

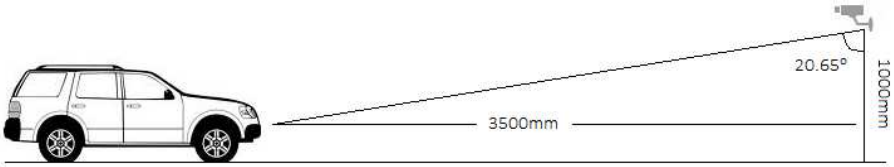

Fig. 6. Camera location

4 Result and Discussion

The program was tested with 1469 real Kazakhstan vehicle photos. The vehicles were taken from different sides and in different climate conditions. The (Fig. 7) illustrates the test cases that were used in testing stage.

Fig. 7. Test cases that were used in testing stage

4.1 Plate Area Detection

The whole tests were divided to five sub-parts. Some sub-parts determines from which side (front or rear) the photo was taken. The Table 1 shows the result of plate area detection using these test cases:

Table 1. Table to test captions and labels

Front	Rear	Sub-part 1	Sub-part 2	Sub-part 3
95%	93.15%	95.5%	94.46%	93.88%

Taking the average value of results above, we will get the whole performance of plate area detection system: (95.3% + 93.15% + 95.5% + 94.46% + 93.88%) / 5 = 94.458%.

4.2 Segmentation

The Kazakhstan number plates were grouped based on their formats in order to segment. The most popular formats are:

- (KZ) DDD LLL DD (8 character)
- (KZ) DDD LL DD (7 character)
- L DDD LLL (7 character)
- L DDD LL (6 character)

Where D is digit and L is letter. The number plates above are new and old types of number plate in Kazakhstan, where (KZ) is a prefix and last two digits identify the region of Kazakhstan.

At the same token, we encounter with the some company cars that have their own number plate formats like L DDDDDD (7 character). Looking to the number plate formats above, we came to conclusion to decompose all number plates based on number of characters. In other words, we have three groups, where in first group - 6 characters (60 number plates), second group - 7 characters (976 number plates) and last group contains the number plates with 8 characters (410 number plates). The result of segmentation shows that total performance of segmentation algorithm is approximately 75%.

4.3 Optical Character Recognition

Some characters like 5 and S, 0 and O and D, 2 and Z, 8 and B are looking similar. Taking into account this fact we test our OCR solution. The performance of our solution is 94%. The detailed explanation of OCR presented in [14].

5 Conclusion and Future Work

Some characters like 5 and S, 0 and O and D, 2 and Z, 8 and B are looking similar. Taking into account this fact we test our OCR solution.

References

1. Kuo-Ming, H., Ching-Tang, H. (2010). A Real-Time Mobile Vehicle License Plate Detection and Recognition. Retrieved from Tamkang Journal of Science and Engineering, Vol. 13, No. 4, pp. 433-442 (2010).
2. Benjapa, R., Kittawee, K., Paruhat P., Thaweesak Y. (2012). License Plate Detection Based on Template Matching Algorithm. Retrieved from International Conference on Computer and Communication Technologies (ICCCT'2012) May 26-27, 2012 Phuket.
3. Beverly, S., Will, H., Peter, L., Patrico, R. Automatic Number Plate Recognition. Retrieved from CS 175 Fall 12.

4. Ondrej, M. (2007), Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems. Retrieved from B.Sc. Thesis, Bruno university of Technology.
5. Yoshimori, S., Mitsukura, Y., Fukumi, M., and N. Akamatsu. (2003) License plate detection using hereditary threshold determine method,Retrived from Lecture Notes in Articial Intelligence, vol. 2773, V. Palade, R. J. Howlett, and L. C. Jain, Eds. New York: Springer-Verlag, pp. 585593.
6. Duan, T., Hong Du, T., Phuoc, T., Hoang, N. (2005). Building an automatic vehicle license plate recognition system,Retrived from Proc. Int. Conf. Comput. Sci. RIVF, pp. 5963
7. Baggio, D., Emami, S., Escriva, D., Ievgen, K., Mahmood, N., Saragih, J. Shilkrot, R. (2012)Mastering OpenCV with Practical Computer Vision Projects. Retrieved from Packt Publishing.
8. Wenjing, J., Xiangjian, H., Huaifeng, Z., Qiang, W. (2007). Combining Edge and Colour Information for Number Plate Detection Retrieved from Proceeding of Image and Vision Computing, pp. 227-232.
9. Reza Azad, Mohammad Baghdadi. (2013). Novel and Fast Algorithm for Extracting License Plate Location Based on Edge Analysis.
10. Baggio, D., Emami, S., Escriva, D., Ievgen, K., Mahmood, N., Saragih, J. Shilkrot, R. (2012) Mastering OpenCV with Practical Computer Vision Projects. Retrieved from Packt Publishing.
11. Wenjing, J., Xiangjian, H., Huaifeng, Z., Qiang, W. (2007). Combining Edge and Colour Information for Number Plate Detection Retrieved from Proceeding of Image and Vision Computing, pp. 227-232.
12. Khurram Khurshid, Imran Siddiqi, Claudie Faure, Nicole Vincent. Comparison of Niblack inspired Binarization methods for ancient documents.
13. CCTV information from the CCTV Advisory Center. Retrieved from: `http://www.cctv-information.co.uk/i/An_Introduction_to_ANPR`.
14. B.Amirgaliyev, M.Kairanbay, Ch.Kenshimov, D.Yedilkhan. Development of Automatic Number Plate Recognition System. Computer Modelling and New Technologies 2014 18(3) p.193-197.

Analysis of the Possibilities for Using a Uniform Bragg Gratingin a Tunable Dispersion Compensator

Piotr Kisala[1], Waldemar Wojcik[1], Aliya Kalizhanova[2], Gulzhan Kashaganova[2], Nazym Kussambayeva[2], and Gulbakhar Yussupova[2]

[1] Politechnika Lubelska, Lublin, Poland
[2] K.I. Satpayev Kazakh National Technical University, Almaty, Kazakhstan
{p.kisala,waldemar.wojcik}@pollub.pl,\\kalizhanova_aliya,guljan_k70,kussambayeva,gulbahar68}@mail.ru

Abstract. The article presents a tunable fibre optic dispersion compensator system, consisting of a specially designed cantilever beam and a uniform Bragg grating. It analyses the group delay and dispersion characteristics in the case that there is no apodization of the grating and also for a grating with apodization used for modulation of the refractive index. Various apodization parameters were tested, along with their effects on the dispersion characteristics of the entire system properties. It is demonstrated in the paper that the apodization parameter affects the compensator group delay characteristic. The finite elements method was used to design a compensator of such a shape that enabled chirp to be induced in a grating of a specified shape. A new design is presented for the system, in which the dispersion properties are tuned by the maximum value of the heterogeneous deformation of the compensator. The paper also includes results showing the effect of the maximum value of heterogeneous stress of the grating on the dispersion characteristics of the proposed construction.

Keywords: dispersion compensation, group delay, fiber Bragg gratings, tunable dispersion compensator.

Introduction

Communication by means of an optic medium has many advantages, but the main factor limiting its efficiency is dispersion. There are many methods of dispersion compensation. The dispersion compensation fibres in use have a flaw in the form of the compensator large size [1]. Another disadvantage is the constant length of compensating fibre required for compensation in a fibre optic connection of a specified length, which means that these solutions cannot be applied in networks with optical switching, where the length of the optical connection may vary. Coupled two-cavity allpass filters realised entirely as thin-film structures for dispersion slope compensation of optical fibres have been proposed [2]. Liquid crystals used in optical sensors [3,4] have also been proposed for the polarization mode dispersion compensation in the Liquid-Crystal Modulator Arrays [5]. Among systems currently used, those based on fibre Bragg gratings are also worth mentioning [6-8]. The basic flaw of such systems is the fact that the grating of the compensator only works for one wavelength. A solution may be to tune the grating in order to adapt its compensatory properties to a specific wavelength for other requirements, e.g. to obtain a defined scope of group delay [9,10]. Chirped [11], non-linear chirped [12], apodized [13] or sampled [14,15] Bragg gratings are most commonly used in the case of dispersion compensators. The influence of the grating parameters on its third-order dispersion value have also been studied [16]. In some cases the acoustic effect allows control of the dispersion using a uniform FBG [17]. The present article presents proposals to use a uniform Bragg grating in the construction of a tunable fibre optic dispersion compensator. The influence of the apodization profile of a homogeneous Bragg grating on the group delay of the signal in the area of the grating was measured. A special

mounting has been proposed for the grating on a specially prepared cantilever beam allowing chirping to be induced in the grating by force being applied. The analysed response to the designed compensator enabled the grating apodization parameter to be identified which allows the optimum spectral characteristics and dispersal properties to be obtained. The first laboratory experiments and simulation tests were conducted using coupled mode theory, and the effect of the parameter of Gaussian distribution of the grating apodization function on its dispersion characteristics. The optical qualities of the system analysed are given as graphs of the group delay and the reflection and transmission spectrum of the gratings in the wavelength function.

Model of a dispersion compensator

In the case of a homogeneous grating, the modulation of the refractive index is approximately homogeneous in the entire grating located within the fibre core, and there are no propagation modes outside the core. With this in mind, plane modes can be disregarded. The distribution of the electric field along the fibre core can thus be expressed as follows:

$$E(x,y,z) = [A(z)exp(-i\beta z) + B(z)exp(-\beta z)]e_t(x,y) \ , \tag{1}$$

where $A(z)$ and $B(z)$are the amplitudes of the propagation modes in directions $+z$ and z respectively. The value z denotes the axis along which the Bragg grating is written, β is the propagation constant. Introducing the distribution $E(x,y,z)$ to the coupled mode equations [18] results in the following:

$$dR(z)/dz = i\sigma(z)R(z) + i\kappa(z)S(z) \ , \tag{2}$$

$$dS(z)/dz = -i\sigma(z)S(z) + i\kappa*(z)R(z) \ , \tag{3}$$

where:

$$R(z) = A(z)exp[i(\delta z - \varphi/2)] \ , \tag{4}$$

and

$$S(z) = B(z)exp[-i(\delta z + \varphi/2)] \ . \tag{5}$$

Here $R(z)$ represents the mode moving in direction $+z$, and $S(z)$ the mode moving in the opposite direction, i.e. z, σ is the direct component (DC) of the modulation of the refractive index, $\kappa(z)$ is the alternate component (AC) of the modulation of the refractive index in the grating (also known as local grating strength). Equations (2) and (3) were used to construct the model of the Bragg dispersion compensator in order to obtain the dispersion and group delay characteristics in the grating. These parameters are also paramount when calculating the parameters of the proposed system.

In the paper, the refractive index of the light in the grating is assumed as a value dependent on z, we can thus say that:

$$n(x,y,z) = n(z) = n_0 + \delta n_0 + \delta n(z)cos(2\pi z/\Lambda + \varphi(z)) \ , \tag{6}$$

where n_0 is the refractive index of the light in the core on the section of fibre without the grating, δn_0 is the depth of modulation of the refractive index, δn_z is the amplitude of the modulation of refractive index, Λ is the grating period, φ is the grating chirp (meaning

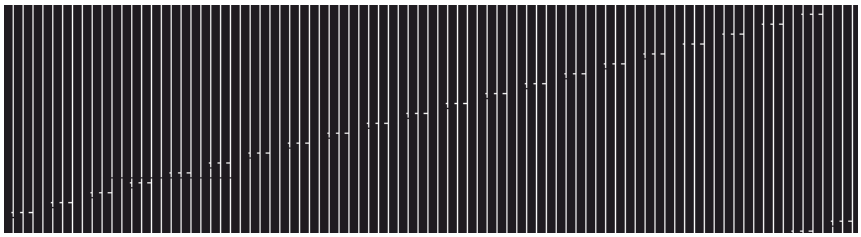

Fig. 1. Designation of the characteristic parameters of the modulation of the refractive index in the model dispersion compensator.

the dependence of the phase of the grating on the position along axis z). The values from the dependencies described by equation (6) are also presented in figure 1.

The value n_1 represents the minimal value of the gratings refractive index, while n_2 denotes the maximum value of the gratings refractive index and L denotes its length. The paper assumes a grating period described as a third degree polynomial:

$$\Lambda = \Lambda_0(b_1 \cdot z^3 + b_2 \cdot z^2 + b_3 \cdot z + b_4) \ . \tag{7}$$

Figure 2 presents the distribution of the changes of the period along the length of the grating.

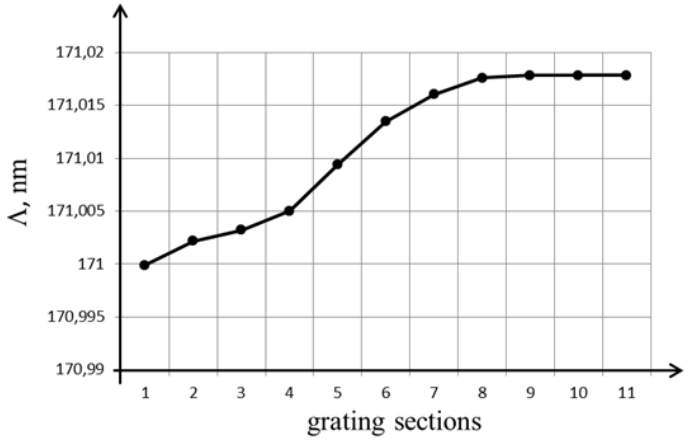

Fig. 2. Change in the period length of the grating used in the dispersion compensator.

For the purposes of solving the coupled mode equations the self-coupling coefficient σ is assumed as equal to:

$$\sigma = \delta + \overline{\delta n_{eff}} 2\pi/\lambda - 0.5 d\phi/dz \ , \tag{8}$$

where δ is the parameter defining the amount of detuning, $\overline{\delta n_{eff}}$ is the *de* index change and is averaged over the extent of each period of the grating, while $d\phi/dz$ defines the grating chirp, connected with the change induced in its periods, according to fig. 2. The detuning parameter was defined based on the following relationship:

$$\delta = \beta - \pi/\lambda = \beta - \beta_B = 2\pi n_{eff}(1/\lambda - 1/\lambda_B) \ , \tag{9}$$

where β denotes the propagation constant, while the Bragg wavelength is expressed as the relationship:

$$\lambda_B = 2n_{eff}\Lambda \,, \tag{10}$$

where n_{eff} is the effective refractive index in the fibre core.

The calculations also assume that the coupling coefficient $\kappa(z)$ can be represented by:

$$\kappa = (\pi/\lambda)\delta n(z)g(z)\nu \,, \tag{11}$$

where $g(z)$ is the function of the apodization, and ν is fringe visibility.

The gratings response in the form of group delay is determined using the state-transition matrix method, with dimensions 2x2. The grating was divided into ten sections, each of which was represented by means of a separate state-transition matrix. The process of propagating light through the entire grating is described by equation (12).

$$\begin{bmatrix} A_{in} \\ B_{out} \end{bmatrix} = \begin{bmatrix} \frac{1}{t} & \frac{r^*}{t^*} \\ \frac{r}{t} & \frac{1}{t^*} \end{bmatrix} \begin{bmatrix} A_{out} \\ B_{in} \end{bmatrix} \,, \tag{12}$$

where $r = |r|exp(i\phi_r)$ and $t = |t|exp(i\phi_t)$ correspond to the reflection and transmission coefficients, ϕ_r and ϕ_t are the phases of the FBG reflection and transmission coefficients. The denotations from equation (12) are presented in figure 3:

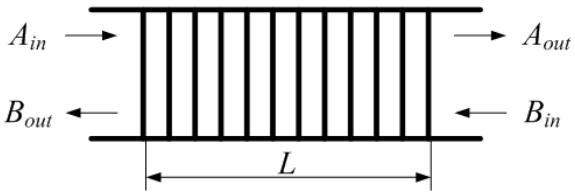

Fig. 3. Denotations of signals in the dispersion compensator.

Thus the whole state-transition matrix, taking into consideration the matrices of the individual sections, looks like this:

$$\begin{bmatrix} {A_{in}}^i \\ {B_{out}}^i \end{bmatrix} = T^i \times \begin{bmatrix} {A_{out}}^{i-1} \\ {B_{in}}^{i-1} \end{bmatrix} = \begin{bmatrix} {T_{11}}^i \, {T_{12}}^i \\ {T_{21}}^i \, {T_{22}}^i \end{bmatrix} \times \begin{bmatrix} {A_{out}}^{i-1} \\ {B_{in}}^{i-1} \end{bmatrix} \,, \tag{13}$$

where index i denotes the number of the grating section, while the individual matrices are expressed as follows:

$${T_{11}}^i = cosh(\gamma z^i) - i(\sigma/\gamma)sinh(\gamma z^i) \,, \tag{14}$$

$${T_{21}}^i = i(\kappa/\gamma)sinh(\gamma z^i) \,, \tag{15}$$

$${T_{22}}^i = {T_{11}^*}^i \,, \tag{16}$$

$${T_{12}}^i = {T_{21}^*}^i \,, \tag{17}$$

while

$$\begin{cases} \gamma = \sqrt{\kappa^2 - \sigma^2} & \kappa^2 > \sigma^2 \\ \gamma = i\sqrt{\sigma^2 - \kappa^2} & \kappa^2 < \sigma^2 \end{cases} . \tag{18}$$

Taking into account the fringe conditions, i.e. $A_{in} = 1$ and $B_{in} = 0$, from equation (19) the gratings responses were denoted in the form of the transmission spectrum A_{out} and reflection spectrum B_{out}:

$$\begin{bmatrix} A_{in} \\ B_{out} \end{bmatrix} = T_N \cdot T_{N-1} \cdot ... \cdot T_i \cdot ... \cdot T_1 \times \begin{bmatrix} A_{out} \\ B_{in} \end{bmatrix} . \tag{19}$$

The paper analyses the work of a dispersion compensator with a non-apodized uniform grating, and with a grating apodized with a Gauss profile, with the entire profile being described by the relationship:

$$g(z) = exp\left[-a((z - L/2)/L)^2\right] , \tag{20}$$

where L is the grating length, $z \in [0, L]$, a is the Gauss parameter.

Method of chirp generation in the dispersion compensator

In order to generate a linearly variable period for the Bragg grating which causes chirp to occur in the grating, the grating was stuck onto a specially designed cantilever beam. This cantilever beam was then subjected to tensile stress which resulted in deformations appearing in the grating. The shape of the cantilever beam and the point where the dispersion compensator grating was attached are shown in figure 4. As can be seen in figure 5, one end (marked B) was placed immobile, while the second (marked A) was subjected to the influence of force which caused it to become deformed. The nature of the deformations in the fibre optic fibre on which the Bragg grating was written is illustrated in figure 6.

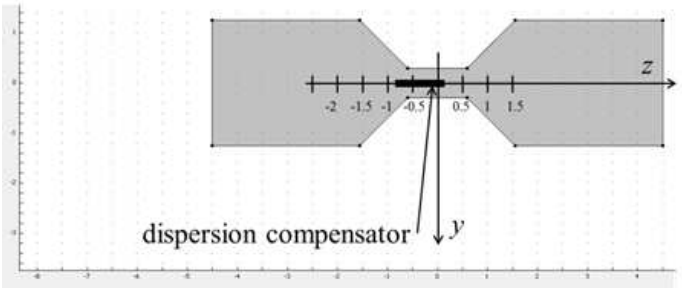

Fig. 4. Shape of the dispersion compensator head, with the position of the Bragg grating marked.

The deformation values were calculated using the finite elements method. As can be seen, the character of the deformation curve corresponds to the polynomial described by equation (7). In the case presented in figure 6, the deformation curve is described by the equation:

$$\varepsilon = c_1 \cdot z^3 + c_2 \cdot z^2 + c_3 \cdot z + b_4 , \tag{21}$$

where the values of the individual constants are $c_1 = 0.002, c_2 = 0.0293, c_3 = -0.0232, c_4 = 1.252$.

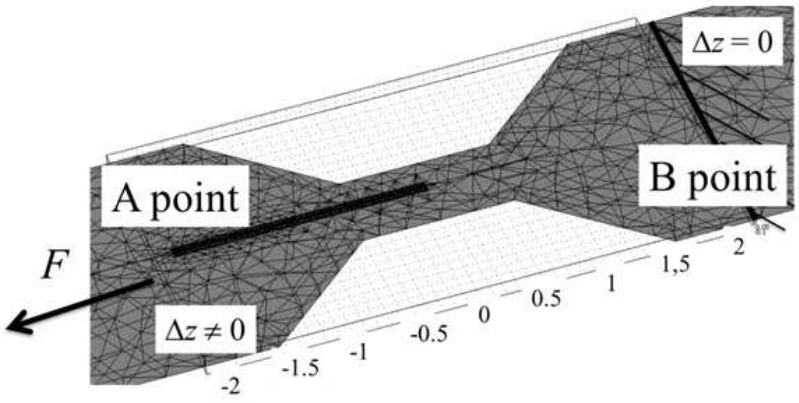

Fig. 5. Grating of finished elements, applied to the dispersion compensator system consisting of a cantilever beam and Bragg grating. End A: subjected to tensile stress F, End B: firmly attached, immovable.

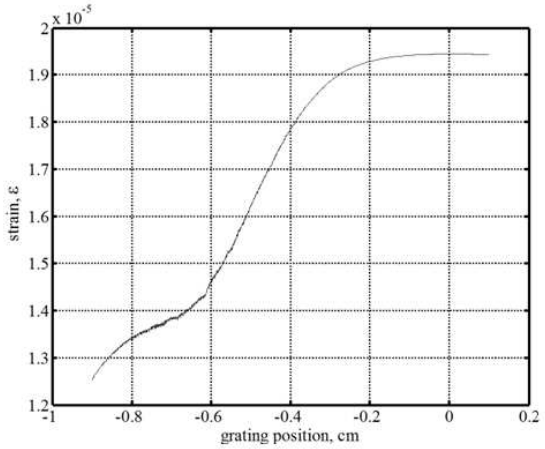

Fig. 6. Distribution of deformation values along the grating. −0.9 beginning of grating, +0.1 end of grating.

Results

In order to establish the dispersion properties of the proposed system, the group delay and dispersion values were calculated. The calculations were carried out assuming the real values of the parameters of the homogeneous grating used as a compensator. A grating with a length of $L = 5$ cm and effective refractive index of $n_{eff} = 1.447$ was used. The Bragg grating was mounted on an element which induced chirp in it with constant force. Its Bragg wavelength was $\lambda_B = 1554.25$ nm. Figure 7 shows the grating's transmission and reflection spectrum, while there is a graph of the group delay in figure 8.

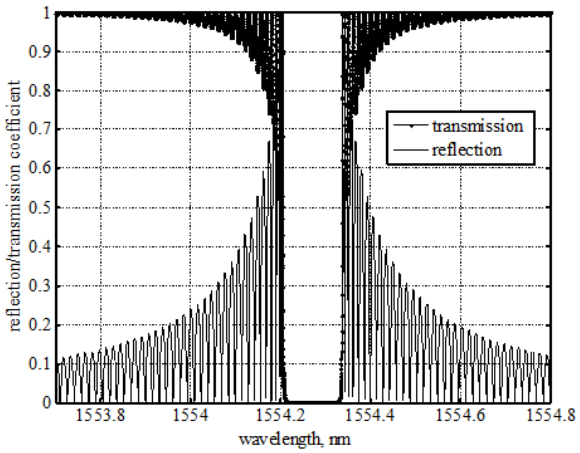

Fig. 7. Spectral characteristics of transmission and reflection in the case of a non-apodized grating, no grating chirp is caused.

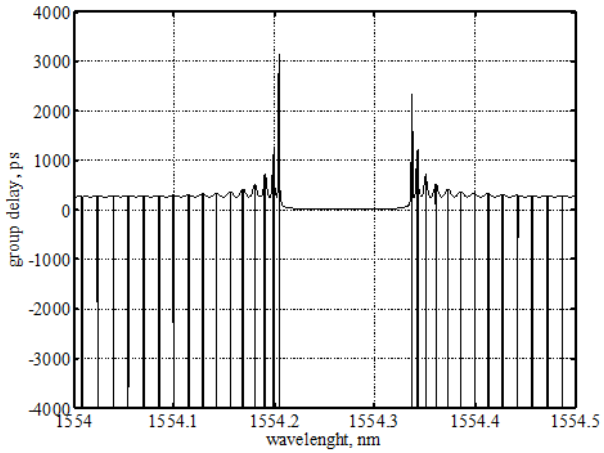

Fig. 8. Group delay where there is no apodization or grating chirp.

Figure 9 presents the dispersion characteristic where there is no apodization or force causing a heterogeneous period to be generated on the length of the grating. Analogous graphs of the

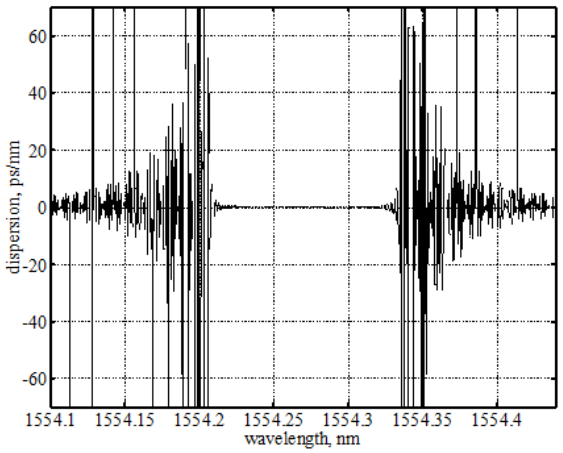

Fig. 9. Dispersion in the area of the wavelength of the compensator Bragg grating with no apodization or grating chirp.

transmission spectrum and of group delay and dispersion in the event that the grating is apodized are given in figures 10-12.

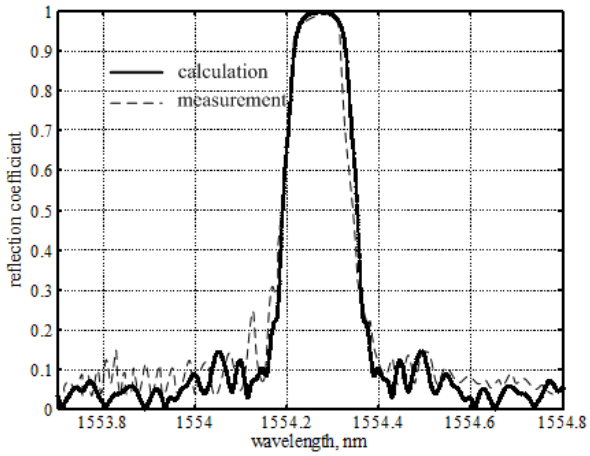

Fig. 10. Measured and calculated transmission spectrum where there is no grating chirp when the Gaussian apodization function ($a = 10$) is applied.

The illustrations above show the relationship between group delay and dispersion for various values of the maximum value of the heterogeneous deformation to which the proposed dispersion compensator system was subjected. An analysis was made of the relationship of the dispersion characteristics depending on the apodization profile of the Bragg grating. The spectral characteristics of the system transmission and reflection were compiled in order to demonstrate the change and shift in the spectrum width, which also affected the system dispersion properties. As can be seen, the increase in amplitude of the deformation causes a broadening of the reflective and transmission spectral characteristics of the compensator. On the basis of the characteristics displayed in figures 9, 11, 13, 15 and 17, a clear decrease can be seen in the dispersion values in the range of the Bragg wavelengths. The fields in which the lowest dispersion values appear

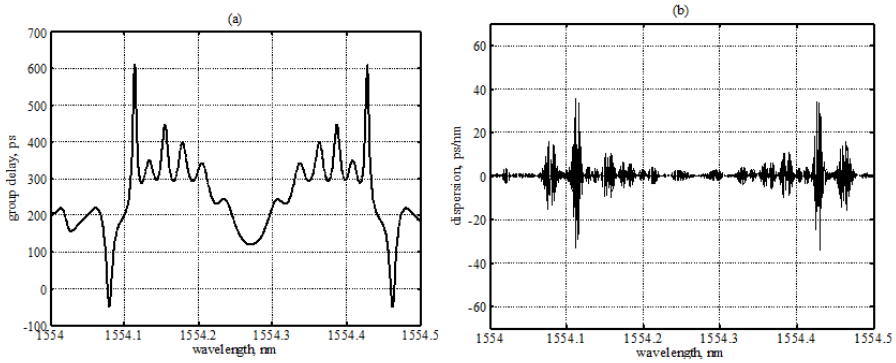

Fig. 11. Characteristics obtained without grating chirp when the Gaussian apodization function ($a = 10$) is applied. (a) group delay, (b) dispersion.

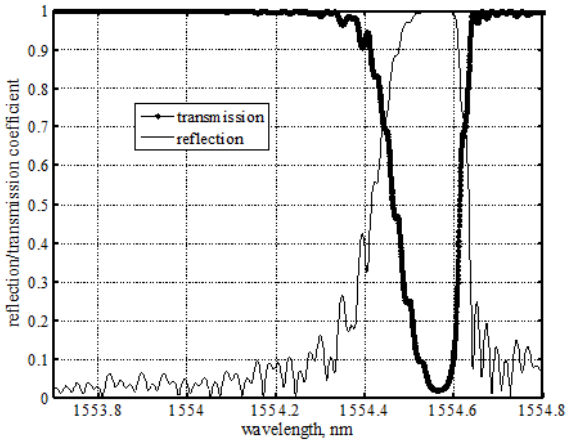

Fig. 12. Transmission and reflection spectrum in the case of heterogeneous elongation of the compensator grating. The maximum value of the grating deformation $= 1.95 \cdot 10^{-5}\varepsilon$. The grating is apodized according to the Gaussian function ($a = 5$).

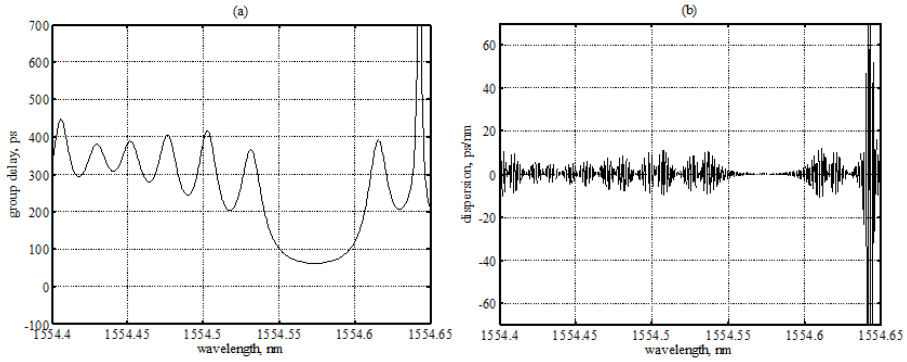

Fig. 13. Characteristics obtained in the case of heterogeneous elongation of the compensator grating. The maximum value of the grating deformation $= 1.95 \cdot 10^{-5}\varepsilon$. The grating is apodized according to the Gaussian function ($a = 5$).

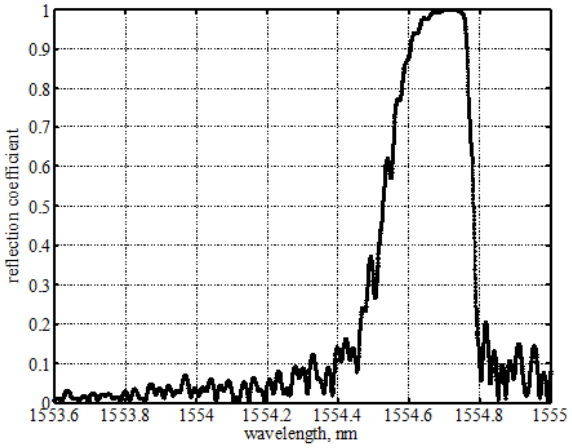

Fig. 14. Reflection spectrum in the case of heterogeneous elongation of the compensator grating. The maximum value of the grating deformation $= 2.92 \cdot 10^{-5}\varepsilon$. The grating is apodized according to the Gaussian function ($a = 5$).

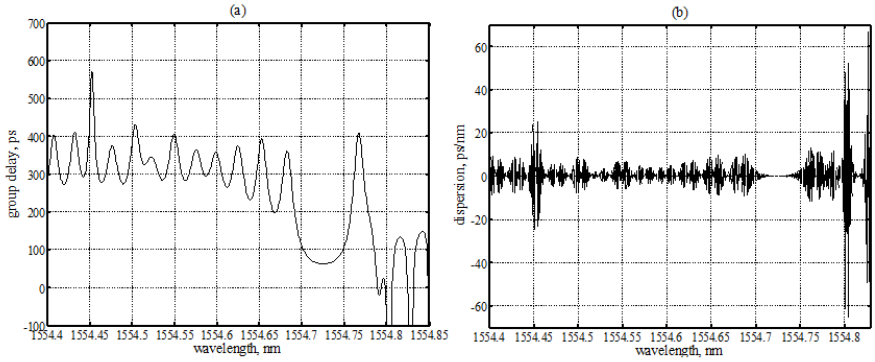

Fig. 15. Characteristics obtained in the case of heterogeneous elongation of the compensator grating. The maximum value of the grating deformation $= 2.92 \cdot 10^{-5}\varepsilon$. The grating is apodized according to the Gaussian function ($a = 5$).

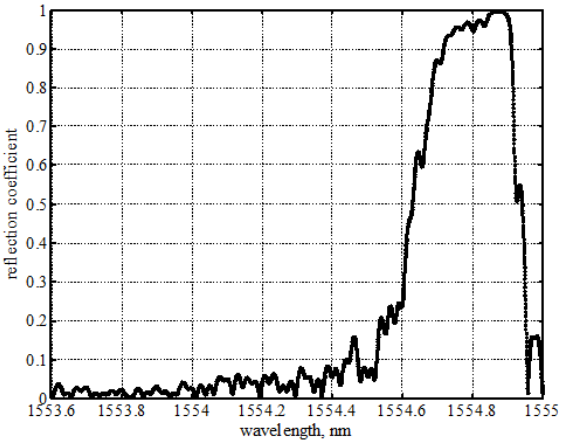

Fig. 16. Reflection spectrum obtained in the case of heterogeneous elongation of the compensator grating. The maximum value of the grating deformation $3.9 \cdot 10^{-5}\varepsilon$. The grating is apodized according to the Gaussian function ($a = 5$).

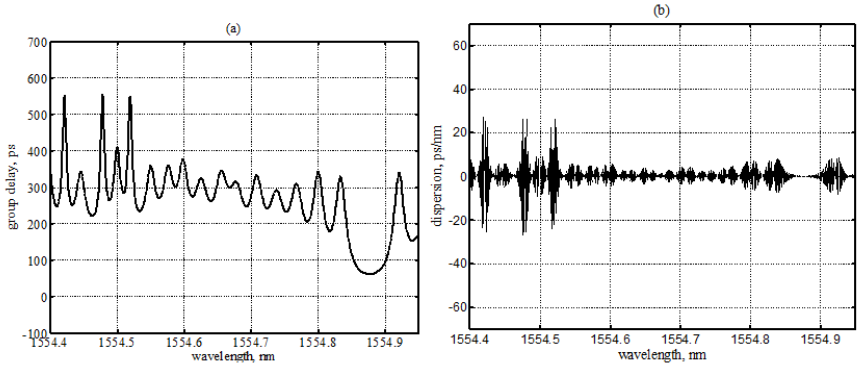

Fig. 17. Characteristics obtained in the case of heterogeneous elongation of the compensator grating. The maximum value of the grating deformation $3.9 \cdot 10^{-5}\varepsilon$. The grating is apodized according to the Gaussian function ($a = 5$).

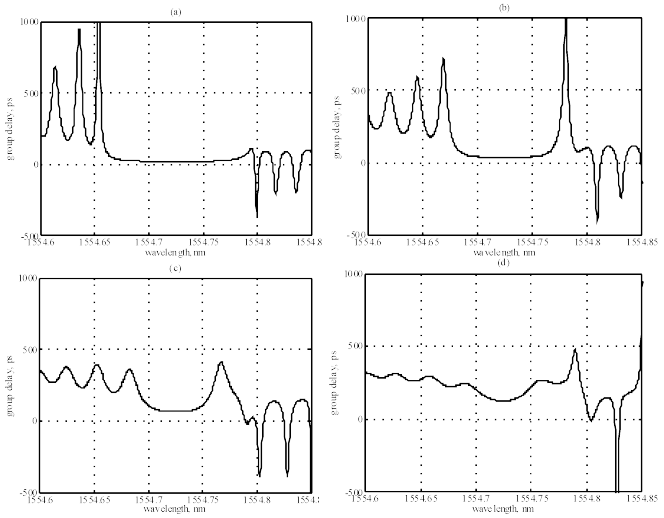

Fig. 18. Graphs of group delay in the vicinity of the dispersion compensator Bragg wavelength, (a) no apodization, (b) Gauss apodization, $a = 2$, (c) Gauss apodization, $a = 5$, (b) Gauss apodization, $a = 10$.

coincide with the lengths of waves which meet the Bragg condition, which are in turn illustrated in figures 7, 10, 12, 14 and 16. The results collated in figure 18 demonstrated that it is possible to control the gradient of the group delay curve by simply regulating parameter a of the given Gauss distribution by relationship (20). It can be seen that along with the increase in the parameter from 0 to 10, this band is reduced and the gradient of the group delay curve decreases, causing a weakening of the dispersion compensation effect. The dynamics of the change in group delay for the non-apodized grating measures 500 ps, while in the case of Gaussian apodization with the parameter $a = 10$ this decreases to 300 ps. The characteristics from figure 18 also show that along with the increase in parameter a of the grating apodization function, the group delay characteristic is smoothed out.

Conclusion

A dispersion compensator was presented which is tuned by the heterogeneous stress of a Bragg grating. Analysis and simulation tests demonstrated that the level of the maximum value of the heterogeneous stress of the grating affects the group delay characteristics and dispersion. It can also be noted that the use of apodization affects the size of the area where the group delay characteristic is flat. By using the negative gradient properties of the group delay characteristic of the proposed system it is possible to apply it in accumulated dispersion compensation in fast data transmission systems. The additionally presented results prove that along with an increase in the maximum value of the heterogeneous deformation of the grating there is an increase in the scope of the wavelengths for which the value of the group delay is a falling curve, there is a simultaneous shift in the wavelengths for which a significant limiting of dispersion takes place. This change is caused by a shifting of the wavelength of the Bragg grating, which is subjected to mechanical stretching. The paper does not analyse the influence of temperature on the dispersion characteristics obtained. In cases where a Bragg grating is used as a dispersion compensator, the influence of the temperature is highly significant, so developing a method for controlling the temperature or eliminating its influence would seem to be of major importance.

References

1. Tewari, R., Basu, M., and Acharya, H.N.: Studying the effect of a central dip on the performance of a dispersion compensated fibre. In: Optics Communications, Vol. 174, Issues 5-6, pp. 405–411 (2000)
2. Jablonski, M., Takushima, Y., Kikuchi, K., Tanaka, Y., Furuki, K., Sato, K., and Higashi, N.: Layered optical thin-film allpass dispersion equaliser for compensation of dispersion slope of optical fibres. In: Electronics Letters, Vol. 36, Issue 13, pp. 1139–1141 (2000)
3. Gotra, Z., Gotra, O., Mikityuk, Z., Zayats, R., and Stadnyc, B.: Fibre-optic sensors on the base of liquid crystals. In: Lightguides and Their Applications, proceedings of SPIE, Vol. 4239, pp. 76–81 (2000)
4. Mikityuk, Z., Lopatynskiy, I., Gotra, O., Dalanbalar, B. Svechnikov, S.V., Kojemiako, V.P., and Kostyukevych, S.A.: Liqud crystal using in optical sensors. In: Selected papers from the international conference on optoelectronic information technologies, proceedings of SPIE, Vol. 4425, pp. 472–477 (2001)
5. Akbulut, M., Weiner, A.M., and Miller, P.J.: Broadband All-Order Polarization Mode Dispersion Compensation Using Liquid-Crystal Modulator Arrays. In: Journal of Lightwave Technology, Vol. 24, Issue 1, pp. 251–261 (2006)
6. Yin, Y., Chen, X., Zhu, L., Xu, K., Lin, Z., Li, X., and Xie, S.: Analytical expression and system simulation of eye opening penalty for chirped sampled Bragg grating dispersion compensator. In: Optics Communications, Vol. 203, Issues 1-2, pp. 93–99 (2002)
7. Ngo, N.Q., Li, S.Y., Zheng, R.T., Tjin, S.C., and Shum, P.: Electrically Tunable Dispersion Compensator With Fixed Center Wavelength Using Fiber Bragg Grating. In: Journal of Lightwave Technology, Vol. 21, Issue 6, pp. 1568–1575 (2003)
8. Tan, Z., Yong, C., Liu, Y., Zhi, T., Jihong, C., Tigang, N., Kai, Z., Ting, C., and Jian, S.: Cross-phase modulation in long-haul systems with chirped fiber Bragg gratings-based dispersion compensators. In: Optik, Vol. 118, pp. 216–220 (2007)

9. Ning, G., Aditya, S. Shum, P., Gong, Y.D., Dong, X.Y., and Ng, J.H.: A PMD compensator with Hi-Bi chirped FBG free from chromatic dispersion. In: Optics Communications, Vol. 245, pp. 153–157 (2005)
10. Lee, Y.J., Bae, J., Lee, K., Jeong, J.-M., and Lee, S.B.: Tunable Dispersion and Dispersion Slope Compensator Using Strain-Chirped Fiber Bragg Grating. In: IEEE Photonics Technology Letters, Vol. 19, Issue 10, pp. 762–764 (2007)
11. Khan, S.S.A., and Islam, M.S.: Performance evaluation of different apodization profiles of linearly chirped FBG for dispersion compensation. In: Computer and Information Technology (ICCIT), pp. 350–354 (2011)
12. Ferreira, M.F., Pinto, A.N., Andre, P.S., Muga, N.J., Machado, J.E., Nogueira, R.N., Latas, S.V., Sousa, M.H., and Rocha, J.F.: Polarization Mode Dispersion in High-Speed Optical Communication Systems. In: Fiber and Integrated Optics, Vol. 24, Issues 3-4, pp. 261–285 (2005)
13. Kolpakov, S.A., Barmenkov, Yu.O., and Aboites, V.: Asymmetrically Apodized Fiber Bragg Gratings for Applications in Dispersion-Less Fabry-Perot Fiber Cavities. In: Fiber and Integrated Optics, Vol. 29, Issue 6, pp. 466–479 (2010)
14. Li, M., and Li, H.: Reflection equalization of the simultaneous dispersion and dispersion-slope compensator based on a phase-only sampled fiber Bragg grating. In: Optics Express, Vol. 16, Issue 13, pp. 9821–9828 (2008)
15. Karim, F., and Seddiki, O.: Study on raised-cosine sampled chirped fiber Bragg grating for dispersion compensation applications. In: Computers and Simulation in Modern Science, pp. 97–99 (2009)
16. Erro, M.J., Laso, M.A. G., Benito, D., Gard, E.M.J., and Muriel, M.A.: Third-Order Dispersion in Linearly Chirped Bragg Gratings and Its Compensation. In: Fiber and Integrated Optics, Vol. 19, Issue 4, pp. 367–382 (2000)
17. Marques, C.A.F., Oliveira, R.A., Canning, J., Pohl, A.A.P., and Nogueira, R.N.: Controlling the properties of Fiber Bragg Gratings based on the Acousto-Optic modulation. In: EUROCON - International Conference on Computer as a Tool, IEEE, pp. 1–3 (2011)
18. Kashyap, R. Fiber Bragg Gratings. In: Academic Press, New York (1999)

Analysis of a Direct Expansion Solar Assisted Heat Pump Suitable for Comfort Applications

Jayaraj.S*[1] ⋆, Lokesh.P[1], Belyayev Ye[2], and Kaltayev A[2].

[1]Mechanical Engineering Department, National Institute of Technology, Calicut 673601, India
[2]Mechanics and Mathematics Department, Al-Faraby Kazakh National University, Almaty, Kazakhstany

Abstract. In this paper the theoretical and experimental studies carried out on a direct expansion solar assisted heat pump (DX-SAHP) under humid weather conditions is reported. Experimental set up used includes a flat-plate solar collectors acting as an evaporator with refrigerant R22, a hermetically sealed reciprocating type compressor, a air cooled condenser and an electronic expansion valve. A system simulation model based on experimental data is developed to predict the thermal performance of the system. Given the ambient parameters (such as solar insolation and ambient temperature), the system simulation model can yield the performance parameters such as energy performance ratio, power consumption, heating capacity, solar energy input ratio. Comparison between the simulation model results and experimental measurements show that, the model is able to give satisfactory predications. Results had indicated that the system performance is governed strongly by the change of solar radiation and ambient temperature.

Keywords: DX-SAHP, Heating Capacity, Performance Predication.

Introduction

Incremental improvements in existing energy networks will not be adequate to supply this demand in a sustainable way. Finding sufficient supplies of clean energy for the future is one of society's most daunting challenges. Given the finite nature of fossil fuel resources, the increasing pollutant emissions (particularly CO_2) caused by their combustion and the change of the earth's climate, innovative technologies for sustaining ecological heat and power generation, especially in the field of solar heating, are gaining more and more importance. Supported by the increase in costs for fossil energy over the last few years, solar heating systems and the use of environmental thermal energy has become an interesting technology for the heating applications.

Heat pumps are devices that are primarily used for space heating and cooling application. In order to improve the performance of heat pump and reduce the fossil fuel and conventional energy consumption in space heating application, renewable energy sources such as solar energy are used as the source of heat pumps. The idea of combining conventional heat pumps and solar systems has taken many interests. The coefficient performance (COP) of heat pumps improves by increasing of evaporator temperature. Since from last 25 years many researchers reported the performance of the solar assisted heat pump by experimentally and theoretically for water heating, space heating and many other applications.

The concept of DX-SAHP was first proposed by Sporn and Ambrose from in West Virginia in the year 1955. Following their work O′Dell *et al.* [1] reported that 10-15% COP improvement and a general procedure is presented to estimating the seasonal performance refrigerant filled solar collector.Chaturvedi *et al.* [2] reported that COP and collector efficiency of a direct expansion solar assisted heat pump system working with R12 as a refrigerant in unglazed solar collector

⋆ corresponding author: sjayaraj@nitc.ac.in.

ranged from 2 to 3 and from 40 to 70%, respectively. Chaturvedi *et al.* [3] further reported the performance of a R22 refrigerant based solar assisted heat pump water heater using unglazed type solar collector. They reported that the performance of the system can be improved by varying the speed of the compressor according to the ambient condition. Morrison [4] experimentally investigated the thermal performance of solar boosted evaporators and a simulation model in the TRNSYS package is developed for assessing the annual performance. Morgan [5] presented the experimental and simulation of thermal performance of a direct expansion solar assisted heat pump system with R11 as a refrigerant, Average values of COP is found to be 2.5 to 3.5. The thermal performance of refrigerant R134a based solar assisted heat pump studied by Saldo *et al.* [6] and reported that, the averaged values of COP ranged from 4 to 9 and solar collector efficiencies 60 to 85%. Ito *et al.* [7] carried out the theoretical and experimental studies on the thermal performance of a heat pump that used a bare flat-plate collector as the evaporator. Theoretical results agreed well with experimental results presented, with average value of COP is found to be 5.3.Huang *et al.* [8] carried out a modeling and system simulation of an integral type direct expansion solar assisted heat pump water heating (DX-SAHPWH). The modeling and simulation assumed a quasi-steady process for all the components except the storage tank.The simulation results agreed very well with experiment. Hawladar *et al.* [9] has been developed to study the thermal performance of solar assisted heat pump water heating system. To validate the simulation model, experiment were conducted metrological condition of Singapore.The values of COP as high as about 9 and average collector efficiency of 75%. Kuang *et al.* [10] have compared the analytical and experimental on DX-SAHP water heating system in which monthly averaged COP is found to vary between 4 to 6 while collector efficiency varied from 40 to 60%. Omojaro *et al.* [11] reviewed and summarized the direct expansion solar assisted heat pump system based on various application and thermal performance. Mohanraj *et al.* [12] experimentally studied the performance of R22 and its alternative refrigerant mixture R407C-LPG and identified that RM30 (LPG10%+ R407C90%) as an optimum refrigerant composition, which has thermodynamic properties closer to that of R22 across the wide range of operating conditions. The energy performance ratio (EPR) of the mixture (RM30) was found to be lower in the range of 2.2-5% compared to that of R22. Hawlader *et al.* [13] investigated the thermal performance of the solar assisted hat pump system for dryer and water heater application. The values of COP, obtained from simulation and experiment are 7.0 and 6.0, respectively. Li *et al.* [14] experimentally analyzed the performance of a direct expansion solar assisted heat pump water heater based on exergy analysis. Results indicate that exergy losses occur in the compressor and solar collector followed by condenser and expansion valve with total exergy destruction of 0.593 kW.Torres *et al.* [15] experimentally predicated the thermodynamic behavior of direct expansion of solar assisted heat pump using unglazed evaporator and reported that exergy efficiency and COP of their system varies from 0.067-0.14 and 2.56-4.36, respectively. Mohanraj *et al.* [16] presented the suitability of artificial neural network (ANN) to predict the performance of a direct expansion solar assisted heat pump (DXSAHP). The ANN approach has been successfully applied for performance prediction of a DXSAHP at different solar intensities and ambient temperatures.

In this paper, focus is made to in developing the experimental set-up model, which can estimate the thermal performance of a DX-SAHP system. The simulation model can predict thermal performance under different ambient parameters and these results are compared with those obtained from the experiment. The influence of various operating parameters on the thermal performance has been identified.

Experiments

Experimental set up

The experiments in the DX-SAHP were carried out during the winter months of 2014 at National Institute of Technology, Calicut (Latitude of 11.15 ^{0}N, Longitude of 75.49 ^{0}E), India. A direct expansion solar assisted heat pump for space heating experimental set-up was designed and fabricated in this study. The specification of each main component is listed in Table 1. The schematic diagram and photographs of the experimental set up are shown in Fig. 1 and 2, respectively.

A glazed type solar collector with total area 2.0 m^2 (2 x 1 m) was used as heat source device as well as evaporator for the refrigerant Difloro-mono Chloro Methane (R22). The solar collector has 0.8 mm thick copper fins with 10mm length are attached with copper tube. The collector/evaporator surface was selective coated to improve its absorptivity of the incident solar radiation. The tube along with fins is placed directly behind the transparent cover with a layer of air separating it from the cover. Bottom side of the solar collector was insulated with glass wool of 25 mm thick to reduce the convective heat loss from the absorber surfaces of the solar collector. The solar collector was designed to take 2.8 kW of heat from both solar and ambient source. An R22 reciprocating hermetic type compressor with rated input power 1020 W is used. To avoid the overload, an inside overheat protector cut-off switches were connected to the compressor. A forced type air cooled condenser was made up of a coil copper tube (9 mm diameter) with face velocity 4.8 m/s. A liquid receiver and sight glass installed downstream of the condenser, followed by the, a filter-drier and a flow meter which are used to remove the moisture content in the refrigerant and measure the flow rate of the refrigerant respectively. The thermostatic expansion valve (external balance type) regulates refrigerant flow through the solar collector/evaporator. In the experiment, the collector/ evaporator was fixed on a plane facing to south at 20^0 angle (the latitude of 11.15^0N, longitude of 75.49^0E Calicut, India) to horizon.

Data Acquisition System

The pressures, temperatures, velocity and flow rate of working fluid (R22) were measured at the locations shown in Fig 1. Also, the ambient temperature, relative humidity, wind velocity and the incident solar insolation were measured.

Suction and discharge pressure of the compressor are measured using mechanical bourdon type pressure gages and pressure transducer at inlet and outlet of the compressor. Similarly, inlet and outlet of the expansion value, the bourdon type pressure gages and pressure transducer are used to measure the condenser outlet pressure and pressure of the refrigerant after expansion. Suction and discharge temperature of the compressor, outlet temperature of the condenser and temperature of after expansion of the refrigerant were measured with platinum resistance thermometers (RTDs). Another, similar RTDs are used to measure the inlet and outlet air temperature of the condenser. The ambient temperature was measured with the help of a thermometer. Relative humidity and wind velocity were measured with help of the Vantage-Pro weather station. Outlet air velocity of the condenser was measured using vane type anemometer and velocity sensor. A solar Pyranometer was mounted on the surface of the solar collector (evaporator) to measure the instantaneous solar radiation. A turbine flow meter was used to measure the refrigerant flow in system. A digital multi-function single phase power meter was used to measure the power consumptions of the compressor. All above temperature, pressure, velocity, solar insolation and

power consumption were monitored and controlled by a personal computer-based data-acquisition system. The data was recorded at every 5 minutes interval in a data logger, which will be later used for the analysis.

Experimental Procedure

To conduct the experiments on direct expansion solar assisted heat pump, initially the system is flushed with nitrogen gas to remove the moisture and impurities present in the system. Then the system is charged with 1100 g of refrigerant, R22 as per the manufacture's specifications. Temperature and pressure at various location of the system like inlet and outlet of the compressor and evaporator/ solar collector, Ambient temperature, Inlet and outlet of the air temperature passing through the condenser, power consumption of the compressor, solar intensity are recorded automatically in the personal computer with the help of data acquisition system at every 5 minutes of interval. The system is found to attain steady state condition within 15-20 minutes.Recorded data corresponding to the steady state conditions are used for further analysis.

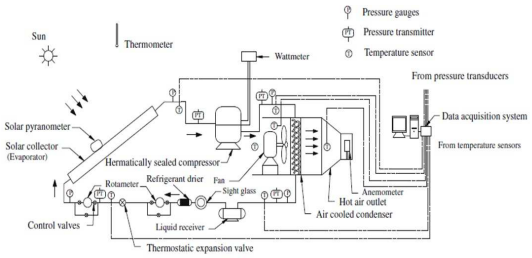

Fig. 1. Schematic view of the experimental set up

Fig. 2. Photographic view of the experimental set up

Model Development

Assumption

A system simulation model is developed to predict the thermal performance of the DX-SAHP system, based on the following assumptions:

- All the process are steady-state within the chosen time interval.

Table 1. Specification of the main components in the experimental Set up

Component	Make/Type	Specification
Compressor	Emerson (M:KCJ515HAE-B422)	Hermetically sealed reciprocating of rated power input 1020 W; Cooling capacity 2750 W; Speed: 2800 rpm; With overload protection.
Condenser	Air Cool (Type : ACD 1520)	Air cooled condenser with 9.52 mm dia; Air velocity through duct:4.8 m/s;
Expansion device	Donfoss India Ltd (Type TX-068Z3206)	Thermostatic expansion valve
Evaporator(Solar Celletor)	Copper Plate	Glazed solar collector of area 2 m^2 with 0.8 mm thick copper fins as absorber surface.
Refrigerant		R22

- Pressure drop is negligible in collector/evaporator, condenser as well as in piping.
- Compression of the refrigerant vapor is assumed to follow a polytropic process.
- Expansion of refrigerant liquid is considered to be isenthalpic.

Based on the above assumptions, governing equations describing the thermal performance of various components of the system have been formulated.

Mass and Energy Balance

In general, the mass balance equation can be expressed in the rate form as:

$$m_i = m_e = m \tag{1}$$

Here, m is the mass flow rate, at inlet and outlet of the system. The mass flow rate of the refrigerant is assumed to be constant at all the typical locations in the system. In this study, the energy balance equation for the DX-SAHP can be expressed as

$$q_c = q_e + P \tag{2}$$

Where q_c is the heat rejected at the condenser, q_e is the useful heat gain at the evaporator /solar collector and P is the work input to the compressor.

Compressor Power Consumption

The performance characteristics of a constant speed compressor will be expressed by their power consumption and cooling capacity. Based on the experimental data the performance characteristics such as power consumption and cooling capacity can be expressed in terms of their evaporating and condensing temperature as follows;

$$P = -17854.2632 + 2381.4993t_e - 73.5881t_e^2 + 870.29110t_c - 9.9902t_c^2 -$$

$$110.8510t_e t_c + 3.4588t_e^2 t_c + 1.2900t_e t_c^2 - 0.0404t_e^2 t_c^2 \quad (3)$$

$$q_e = -76112.1029 + 10754.2169t_e - 350.0243t_e^2 + 3345.3171t_c - 36.9126t_c^2 -$$

$$479.2131t_e t_c + 15.9065t_e^2 t_c + 5.4086t_e t_c^2 - 0.1809t_e^2 t_c^2 \quad (4)$$

Where t_e and t_c are the evaporating and condensing temperatures. The constants applicable to the above equations are determined by a suitable equation fitting procedure.

Evaporator/Solar Collector Model

The useful energy gain of from the flatplate collector q_e, operating at steady state conditions can be evaluated as follows;

$$q_e = A_{cl}(I \pm -U_L(t_e - t_{amb}) \quad (5)$$

Where A_{cl} is the area of the solar collector, I is the incident the solar radiation on the solar collector, U_L is the overall heat loss coefficient from the collector to the ambient air, t_e is the average evaporating temperature of collector/evaporator unit, and t_{amb} is the ambient air temperature.

Condenser Model

Heat ejected at the condenser is the sum of energy absorbed in the evaporator(solar collector) and the work in input in compressor. It also expressed as

$$q_c = m_{air} C_{pa}(t_c - t_{amb}) \quad (6)$$

Where m_{air} is the mass flow rate of air through the duct area, C_{pa} is the specific heat of air, t_c is condensing temperature of condenser unit and t_{amb} is the ambient air temperature.

Energy Performance Ratio

Energy performance ratio (EPR) or the performance coefficient of the system can be expressed as follows;

$$EPR = q_c/P \quad (7)$$

Where q_c is the heat rejected at the condenser and P is the work input to the compressor.

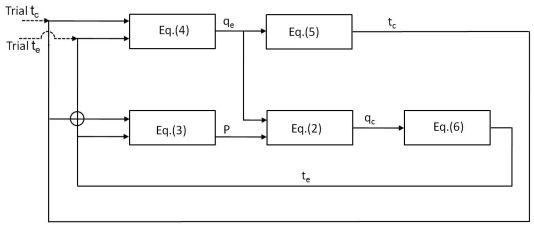

Fig. 3. Information flow diagram of the simulation

Simulation Procedure

Based on the above analysis of each components of the proposed DX-SAHP system, a Matlab® (Version 9.1) program was written and developed to estimate the thermal performance of this system. The input data are ambient parameters (such as solar intensity and ambient temperature). The technique chosen for the mathematical simulation was Newton-Raphson method, which is the most popular technique for solving the simultaneous equation. The information flow diagram of the simulation program carried out is shown in Fig. 3.

Result and Discussion

The experiments were carried out under the metrological condition of Calicut located in the southern peninsula of the Indian continent. To determine the various performance parameters such as power consumption, heating capacity, energy performance ratio, solar energy input ratio and compressor discharge temperature of a DX-SAHP system. The results obtained from the experimentation and system simulation model, at different ambient parameters (such as solar insolation and ambient temperature) are compared and presented in this section.

Experimental Observation

The variation in refrigerant temperature at typical locations in the DX-SAHP system working with R22 as refrigerant is shown in Fig. 4. From figure it is observed that the average temperature of the refrigerant measured at the compressor suction and exit are 15.6 and 71.0^0C, respectively. The temperature at the compressor outlet is the major factor affecting the life of the compressor. Similarly the average temperature of refrigerant at inlet and outlet of the expansion valve are 43.9 and 15.9^0C, respectively. The variation in ambient temperature and air temperature at the condenser outlet are shown in the Fig. 5. During the experimentation, the ambient temperature varies between 29.2 and 33.9^0C with an average temperature of 32.3^0C. The air temperature at the condenser exit varies between 38.4 and 49.9^0C with an average temperature of 45.0^0C. All the observed temperatures at the typical locations in the heat pump circuit and the air temperature at entry and exit of the condenser get increased with increase in solar insolation as expected.

The variation in pressure at various points in the refrigeration system is shown in Fig. 6. As expected, the pressure at compressor suction and compressor discharge varies from 0.47 to 0.537 MPa and from 1.59 to 2.01 MPa, respectively with increase in solar insolation from 0 to 918 W/m^2. Increase in solar insolationwill increase the load on solar collector. Hence, the temperature and pressure at all typical locations in the heat pump circuit and the temperature of the air at entry and exit of the condenser get increased with increase in solar insolation.

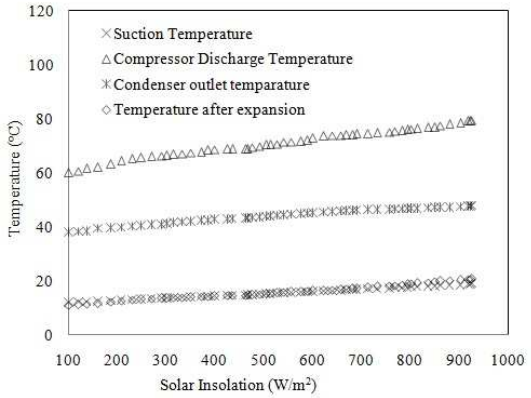

Fig. 4. Variation of temperature at different location in the DX-SAHP system

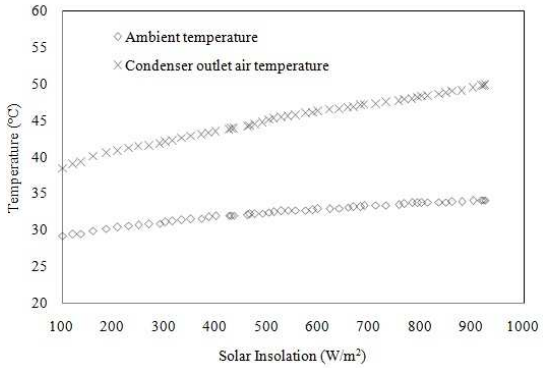

Fig. 5. Variation of ambient and condenser outlet air temperature with respect to solar insolation

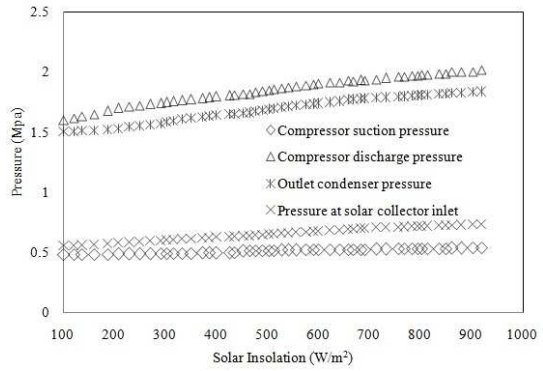

Fig. 6. Variation of pressure at different location in the DX-SAHP system

Comparison of Experimental and Simulation Results

A series of experiment were conducted during the winter months of 2014 under the metrological conditions of Calicut city located in the southern peninsula part of India to reflect the performance of DX-SAHP system during the winter period, and results are presented in this section. Power consumption, heating capacity, energy performance ratio and solar energy input ratio predicated by simulation and the experimental results are compared.

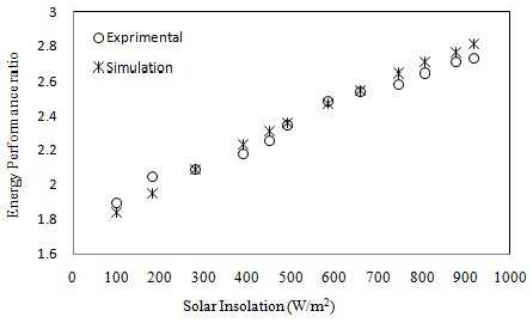

Fig. 7. Energy performance ratio (COP) with respect to solar insolation

The simulation predicted and calculated energy performance ratio values against the solar intensity are reported in Fig. 7. For this parameter, it is seen that, initially actual performance of the system is higher than simulation. It is because the collector was able to gain more latent heat from the solar and ambient, which improved the thermal performance of the system, and values varied from 1.88 to 2.72. Compared to the experimental results, the simulated values of the of the energy performance ratio have same trend with average error 1%. Energy performance ratio is the most important performance parameter considered for rating of a DX-SAHP, which is calculated based on condenser capacity and compressor power consumption.

The power consumption predicated by simulation and experimentally measured results against the solar intensity are compared in Fig. 8. In this case, these results confirmed that the experimental values are closer to simulation predicted results with average error of 2.05%. Compressor power consumption is an important factor considered for rating of heat pump performance.

A plot of simulation predicted and experimental values for the heating capacity of the DX-SAHP with respect to solar intensity are depicted in Fig. 9. The comparison of results showed that simulation predicted heating capacity is found to be closer with experimental heating capacity with average deviation in the range of 2%. Heating capacity is a major factor considered for selection of DX-SAHP for specific applications.

Solar energy input ratio (SEIR) is also an important parameter considered for evaluating the performance of the solar collector used in the DX-SAHP. The SEIR influence the performance of whole system. The SEIR varies between 0.20 and 0.80 with increase in solar intensity from 100 to 918 W/m^2. Simulation predicted and experimental measured values of solar energy input ratio (SEIR) against solar insolation are shown in Fig. 10.

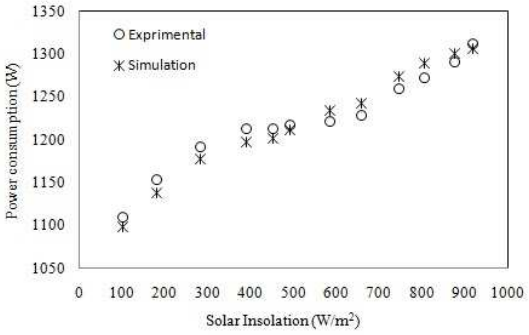

Fig. 8. Power Consumption with respect to solar insolation

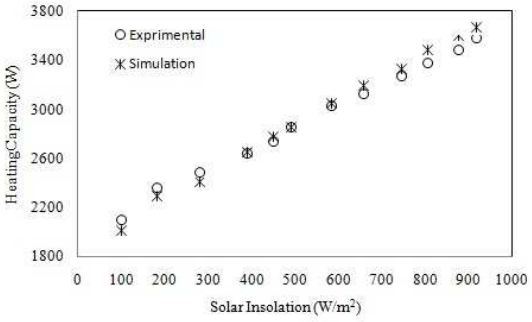

Fig. 9. Heating Capacity with respect to solar insolation

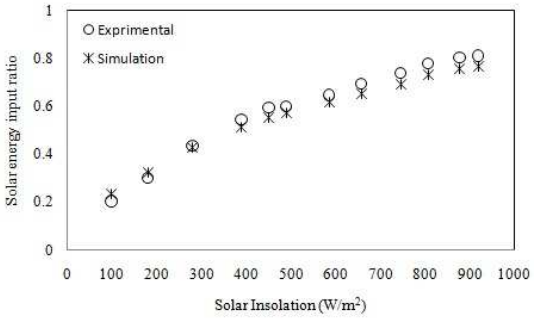

Fig. 10. Solar energy input ratio with respect to solar insolation

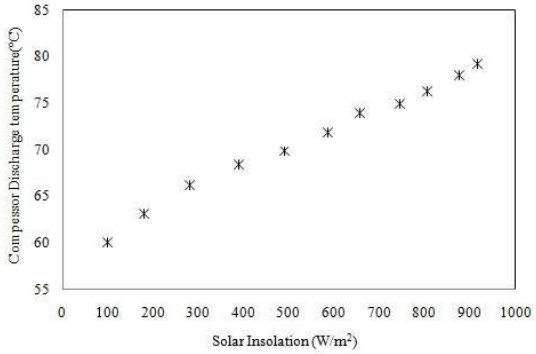

Fig. 11. Compressor discharge with respect to solar insolation

Figure. 11. Shows the compressor discharge temperature of the DX-SAHP corresponding system is compared with different solar insolations. The compressor discharge temperature is major influencing parameter in the DX-SAHP. If the discharge temperature of the refrigerant from the compressor becomes too high, it may result in breakdown of the oil causing excessive wear or reduced life of the valves. The compressor discharge temperature of the refrigerant various between 60.1 to 79.9^0C with increase in solar intensity in the range of 100 to 918W/m^2.

Conclusion

The thermal performance of direct expansion solar assisted heat pump system under metrological condition of Calicut in India, have been investigated and experimental results were obtained. Energy performance ratio is found to be vary from 1.88 to 2.72 with power consumption variation from 1110 to 1310W. The heating capacity range is from 2.09 to 3.57 kW while solar energy input ratio variation from 0.2 and 0.8. These performance parameters results confirmed that, the experimental values were agreed well with simulation predicated results with average error of 1-2%. The system simulation technique presented in this paper is expected to contribute to further studies and applications of DX-SAHP systems with ranging operating conditions and refrigerants in the future.

References

1. O′Dell, M. P., Mitchell, J. W., Beckman, W. A. : Solar heat pump systems with refrigerant filled collectors J. Solar Energy Engineering. 106, 519-525 (1984)
2. Sushil, K., Chaturvedi, James Y., Shen.: Thermal performance of a direct expansion solar-assisted heat pump. J. Solar Energy. 33, 155-162 (1984)
3. Chaturvedi, S.K., Chen, D.T., Kheireddine A.: Thermal performance of a variable capacity direct expansion solar-assisted heat pump. J. Energy Conversion and Management. 39, 181 - 191 (1998)
4. Morrison, G.L.: Simulation of packaged solar heat-pump water heaters. J. Solar Energy. 53, 249-257 (1994)
5. Morgan, R.G.: Solar assisted heat pump. J. Solar Energy. 28, 129-135 (1982)
6. Vladimir Soldo, B., Tonko Curo., Thermal performance of a direct expansion solar assisted heat pump, Refrigeration and Air Conditioning pp. 724-731. (2004)
7. Ito, S., Miura, N., Wang, K.: Performance of a heat pump using direct expansion solar collector. J. Solar Energy. 65 189-196 (1999)
8. Huang, B.J., Chyng, J.P.: Performance characteristics of integral type solar-assisted heat pump. J. Solar Energy. 71, 403-414 2001
9. Hawlader, M.N.A., Chou, S.K., Ullah, M.Z.: The performance of a solar assisted heat pump water heating system J. Applied Thermal Engineering. 21, 1049-1065 (2001)
10. Kuang, Y.H. and Sumathy, K. and Wang, R.Z.: Study on a direct-expansion solar-assisted heat pump water heating system. International Journal of Energy Research. 27, 531-548 (2003)
11. Peter Omojaro., Cornelia Breitkopf.: Direct expansion solar assisted heat pumps: A review of applications and recent research. J. Renewable and Sustainable Energy Reviews. 22, 33-45 (2013)
12. Mohanraj, M., Jayaraj, S., Muraleedharan, C.: Comparison of the performance of a direct expansion solar assisted heat pump working with R22 and a mixture of R407C-LPG. International Journal of Thermal Sciences. 48, 1036-1042 (2009)
13. Hawlader, M.N.A., Chou, S.K., Jahangeer K.A., Rahman S.M.A., Eugene Lau K. W.: Solar-assisted heat-pump dryer and water heater. J. Applied Energy. 74, 185-193 (2003)
14. Li, Y.W., Wang, R.Z., Wu, J.Y., Xu, Y.X.: Experimental performance analysis on a direct-expansion solar-assisted heat pump water heater. J. Applied Thermal Engineering. 27, 2858-2868 (2007)
15. Torres, R.E., Picon Nuez, M., Cervantes de G, J.: Exergy analysis and optimization of a solar-assisted heat pump. J. Energy. 23, 337-344 (1998)
16. Mohanraj, M., Jayaraj, S., Muraleedharan, C.: Performance prediction of a direct expansion solar assisted heat pump using artificial neural networks. J. Applied Energy. 86, 1442-1449 (2009)

Raspberry Pi, Mathematica, and Electrical Engineering Education

Miroslav Lutovac, Petar Spalevic, and Nebojsa Arsic

Singidunum University,
Danijelova 32, 11000 Belgrade, Serbia
{mlutovac}@singidunum.ac.rs
http://singidunum.ac.rs
Faculty of Technical Science Kosovska Mitrovica, University of Pristina,
Kneza Milosa 7, 38220 Kosovska Mitrovica, Serbia
{petar.spalevic,nebojsa.arsic}@pr.ac.rs
http://ftnkm.rs/

Abstract. This paper presents a new approach of teaching electronics courses and in general electrical engineering teaching using a low cost computer with free software and computer algebra systems. Computer algebra system (CAS) is used to formulate the circuit equations and prepare for symbolic solving and thus completely specify an electrical circuit by using Mathematica as CAS. Exercises introduce diodes, transistors, operational amplifiers, and their use in simple real-world circuits. From symbolic expressions it is possible to design and optimize specific circuits. Other software programs such as LTSpiceand LABVIEW can be used for simulating of final solutions and verification. These way students are free from manual solving large system of equations and are focused on understanding the functional model and simulation of complex electrical systems with complex devices.

Keywords: Analog electronics, Digital electronics, Electrical engineering education, Computer algebra system/

1 Introduction

The Raspberry Pi is a low-cost small-size computer [1] that enables people of all ages to explore computing and to learn how to program or use programming environment like Wolfram Mathematica [2]. In recent years, computer education has focused largely on office skills and not on understanding how computers work, or how you can use them to create new programs and inventions. The Raspberry Pi redresses the balance. It can be used for games, music, photo editing, and word processing, like any computer. But it can do so much more, providing a gateway into programming and electronics. You can use simplified operating system on an SD card and connect the device to a screen, mouse, keyboard, and Internet. The new concept of teaching electronics courses to electrical engineering students uses a numeric solvers and computer algebra system (such as Mathematica) and low-cost computers for teaching, analysis, and design electronic devices that can be put in working environment with the Raspberry Pi. The methodology is based on more general approach using symbolic processing [3]. Computer algebra system (CAS) is used to formulate circuit equations and prepare for symbolic solving and thus completely specify an electrical circuit by using Mathematica as CAS. Exercises introduce simple real-world circuits [4]. From the symbolic expressions it is possible to design and optimize specific circuits [5]. These way students are free from manual solving large system of equations and are focused on understanding the functional model and simulation of complex electrical systems with complex devices. The new approach is mainly targeted to combine understandings of simple devices, to integrate models into large systems, and to use computer algebra systems instead of manual derivation of responses [6]. Combining visual interpretation from mathematical models

with simulating numeric tools, electronic kits, and the virtual instrumentation, the process of analysis and design are integrated into unique development environment that can be used by students of electrical engineering.

2 Graphics User Interface using Drawing Software

A graphical user interface (GUI) is a type of interface that allows users to interact with electronic circuits and components through graphical icons placed on a pallet and visual indicators. This is opposed to text-based an interface that is typing command. Most of typing errors on a command line are main reason for delay in development of a new circuit analysis. Unfortunately, some higher level programming environments for using in engineering has no appropriate GUI for solving student or research problems. This can be solved using other tools, such as LTSpice [7] for generating textual description of elements and their connections on an imaginary board, and later to translate into another textual description that can be used for drawing the schematic. CAS such as Mathematica [2] can transform one textual description into another because it has power commands for symbolic manipulation. Similar drawing commands are already available in some Mathematica application packages [8].

2.1 LTSpice

An electronic circuit can be verified without making a real printed circuit board or a soldering iron. Circuit simulation can save costs, time, and efforts when analyze existing circuit, designing a new device, or modifying existing electronic solutions [7]. The SPICE program was developed before 35 at the Berkeley University using FORTRAN programming language. The version for PC computers was called PSPICE. Linear Technologies, one of the leading manufacturers of electronic components, offers a free full SPICE-program named LTSpice without any restrictions. It can be downloaded from the web without any problems or fees but the usage is a little tricky a mixture of command lines, GUI (graphics user interface), and mouse clicks. Fortunately, many tutorials are available for students and engineers. Engineers and students that deal with analog electronics heavily rely on simulation to predict circuit features. The value of a simulator depend on how well it can predict real physical component characteristics, and how quickly it can produce acceptable results. The difference between simulated and real performance can be a reason for low production yield, and very costly iterative debugging procedures. SPICE is used for analog circuit simulation because it can compute the small (linear) and large signal (nonlinear) behavior of arbitrary circuits [9]. Three numerical methods are used in SPICE for analog circuit simulation: (a) Newton iteration to find the solution of circuits with nonlinear elements, (b) Sparse matrix methods to corral huge matrices into the address space of a practical computer, and (c) Implicit integration to integrate the differential equations that arise from circuit reactance. In the transient and steady-state analysis, usually we start by writing circuit equations using the Ohms law, the Kirchhoff s current law (the net current entering a node is zero), and the Kirchhoff s voltage law (the algebraic sum of the voltages for any closed path in an electrical circuit is equal to zero). Because the current voltage relationships for inductances and capacitances involve integrals and derivatives, the equations can be converted to pure differential equations by differentiating with respect to time. The study of transients and steady-state requires solving differential equations, which can be unpleasant for young students of electrical engineering and electronics. The main feature of circuits can be described using existing circuits from text-books [4] or [10], such as an amplifier with a single transistor, presented in Fig. 1.

The analysis can start by redrawing schematic from the book. Most electric components have available manufacturers attributes (resistors, capacitors, inductors, diodes, bipolar transistors,

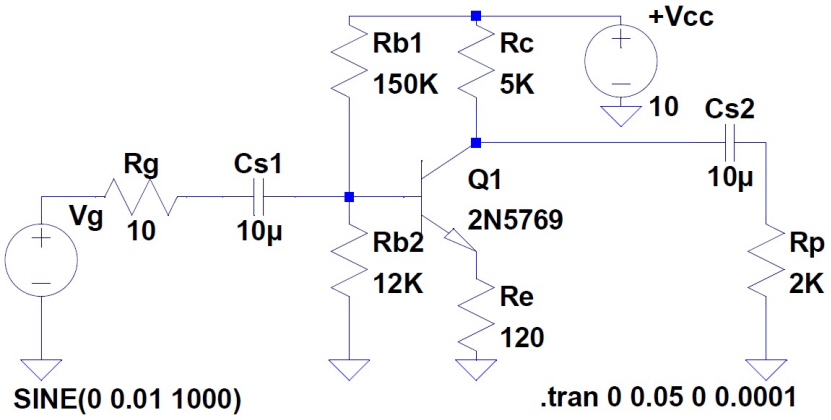

Fig. 1. Transistor amplifier with common emitter and negative feedback

MOSFET, JFET). Before running a simulation, the type of analysis should be defined (small signal AC, DC sweep, noise, DC transfer function, DC operating point). Analysis of an electrical circuit includes the determination of voltages and currents given the element numeric values. The time domain and frequency domain analysis of the circuit from Fig. 1 is presented in Fig. 2 and Fig. 3.

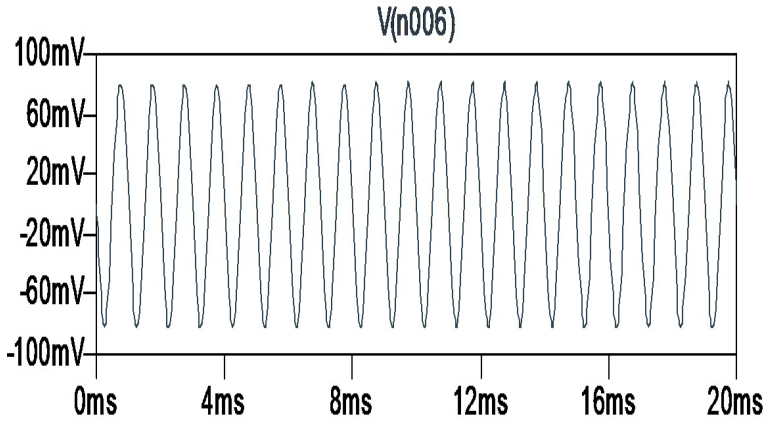

Fig. 2. Time domain simulation

Gain as magnitude and phase response is shown in Fig.3, where unwrap function is used for plotting phase response between 0 and -360^o, (correct phase angles to produce smoother phase plots), instead of response only from -90^o to 90^o. From Fig. 3 it follows that the analyzed circuit has the transfer function poles at very high frequencies although we cannot see capacitance at the Fig. 1. Actually, the parasitic capacitances between base and emitter and base and collector exit in transistor model Q1 2N2222 and thus the simulation is very close to the real measured characteristic of the circuit from Fig. 1.

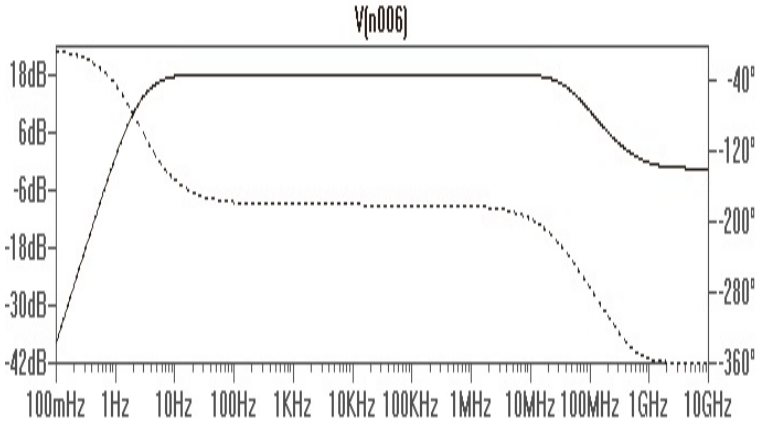

Fig. 3. Frequency domain simulation

2.2 SchematicSolver

SchematicSolver [8] is a powerful and easy-to-use schematic capture, capable for symbolic analysis, processing and implementation of linear and nonlinear systems in Mathematica. With a minimum understanding of basic system theory, students can successfully design and simulate various systems. For freshmens, it is perfect for learning and experimenting with system analysis, implementation and design. For researchers, symbolic analyses and symbolic processing provide a sophisticated environment for testing and trying all the so called "what if"scenarios for system design. Research projects can be accomplished more in less time than with traditional prototyping methods. Fig. 4 presents several palettes and part of drawing board. The graphical representation of a system is not a frozen picture. Once when you have a schematic of some system, and when you find out that the same structure can be used to build a schematic of another, you can do that by simple replacements of any element because the description of a system is a list of strings.

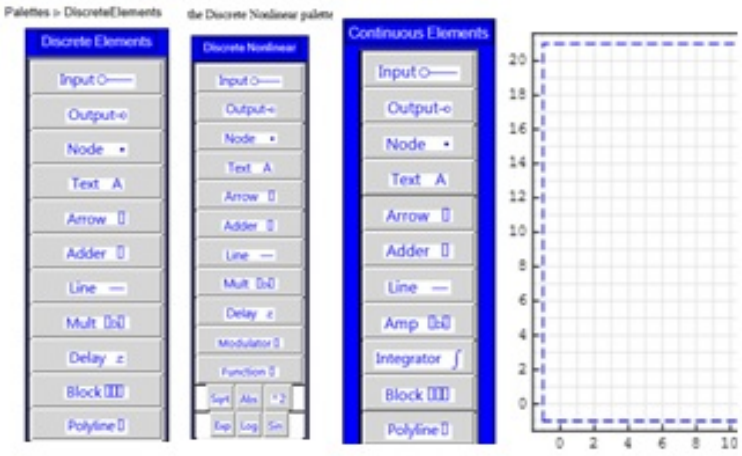

Fig. 4. Palettes and drawing board

Palettes provide an easy point-and-click interface for performing the most common drawing tasks.

3 Raspberry Pi and Mathematica

The Raspberry Pi is a single board computer. It requires additional hardware: prepared operating system on SD card, USB keyboard, USB mouse, display (HDMI input High Definition Multimedia Interface), power supply, internet connectivity a USB WiFi adaptor or a LAN cable. The processor at the heart of the Raspberry Pi system is a system-on-chip (SoC) multimedia processor (including central and graphics processing units, the audio and communications hardware). It operates on the 5V 1A power supply. It is not compatible with traditional PC software. But there is a plenty of software available for the Raspberry Pi. The Raspberry Pi is designed to run an operating system called GNU/Linux referred to simply as Linux that is open source it is possible to download the source code for the entire operating system. Nothing is hidden and all changes are made in full view of the public. Several versions of Linux have been ported to the Raspberry Pi s chip, including Debian, Fedora Remix and Arch Linux. The majority of modern Linux distributions are user-friendly, with a graphical user interface (GUI) that provides an easy way to perform common tasks, exactly what we really need for teaching electrical engineering by experiments. Because of its origins in embedded computing, the Raspberry Pi does not have anything like a PC s BIOS menu where various low-level system settings can be configured. It relies on text files containing configuration strings that are loaded by the chip when the power is switched on. One of tasks for the RPi (Raspberry Pi) is to carry out is that of a home theatre PC (HTPC) because it is specifically designed as a multimedia powerhouse (a convergence device that combines the capabilities of a personal computer with a software application that supports video, photo, audio playback, and video recording functionality). The graphics portion of this system-on-chip, a Broadcom Video Core IV module, is capable of full-speed high-definition video playback using the H.264 format. Although it will never reach the same levels of performance as a standard desktop or laptop, the flexibility of the Raspberry Pi makes it a good choice as a low-power general-purpose desktop computer. Cloud-based software offers powerful using office-centric software on the RPi. Cloud-based software has other advantages. A cloud-based application looks the same on all devices, including mobile-oriented versions designed for access from smartphones and tablets. Files are also stored on the remote servers, making them accessible from any device without taking up any room on the RPi SD card. The RPi can work also as a Web Server. A full version of the Wolfram Language is available for the Raspberry Pi computer and comes bundled with the Raspbian operating system. Programs can be run from a RPi command line or as a background process, as well as through a notebook interface on the RPi or on a remote computer. On the RPi, the Wolfram Language supports direct programmatic access to standard RPi ports and devices, including camera. Some of examples can be found on the web site http://www.wolfram.com/raspberry-pi/ such as Building a GPS tracker with the Raspberry Pi, Using the GPIO with the Wolfram Language + Raspberry Pi, Checking the weather with the Wolfram Language + Raspberry Pi, Snapping pictures with the Wolfram Language on the Raspberry Pi, Reading Temperature Sensors in the Wolfram Language on the RPi, How to build a CCX (Cisco compliant extensions) Wifi Telemetry tag?, Bitcoin tipping machine with Raspberry Pi and chain.com, and many more. The Mathematica (now referred to as Wolfram Language) is a general multi-paradigm programming language that serves as the main interfacing language [2] and the Wolfram Programming Cloud. Main features are symbolic computation, functional programming, and rule-based programming. It can be used for representing arbitrary structures and data.

4 Electrical Engineering Education

Over the past 100 years engineering education has passed several phases with emphasizing different objectives [11]. According to some researches several topics are still challenging. Two phases are still important such emphasis on engineering design and prevalence of information, communication, and computational technologies in engineering education [11]. One reason for this shift in engineering education is that the emphasis on engineering science and mathematics has gone too far. An effort is made to bring the practical side of engineering design back into the engineering curriculum so that students upon graduation can be prepared for industrial practice. Instead of introducing computer programming courses as freshman courses more time student should use computer as computer algebra system for learning mathematics. In a similar way as calculators free students of using slide-rulers 40 years ago for basic calculations, CAS can replace numeric software tools for derivations and proving theorems. Thus, students can spend more time on understanding basic electrical theory and phenomena they are learning instead of improving mathematical skills. The computational power of CAS is impressive: graphing, equation solving, and symbolic manipulation are some of the capabilities available in a handheld device.

5 Design by analysis using CAS

Design of a specific device can start using the circuit configuration given in textbooks and technical journal papers. The concept of abstraction is useful to unify the set of engineering simplifications made in the design. Usually the problem is under-constrained so it has many answers. Interactive designs and numeric simulations allow designers to quickly iterate the proposed design solution and evaluate the merits of different ideas, compare alternatives, and identify design weaknesses before a costly production has been done. A unique approach called design space is introduced in [3]. This approach enables to study the performance of structures by varying design parameters in a cost effective way. This signifies the transition from the traditional design by formula to design by analysis approach. However, there are still many practical issues that have to be addressed to get the real implementation.

5.1 Load-Line Analysis (DC Analysis) using CAS

In the next example we propose symbolic design that provides the voltage at the collector to be half of the voltage supply. The design assumes to find a resistor value from available set of values. After drawing the schematic of the circuit, such as that in Fig. 1, a set of equations can be derived using the Ohm's law, the Kirchhoff's current law, and the Kirchhoff's voltage law, that is programmed in Mathematica [2]. Two unknown variables are the base current and the voltage between collector and emitter nodes, therefore set of two equations is required (set of equations is denoted by **equDC**):

$$\begin{aligned} equDC = \{&Vcc/Rb1 - Ib - (Vbe + Ib(1+b)Re)/Rb2 == 0, \\ &Vcc - bRcIb - Vce - Ib(1+b)Re == 0\} \end{aligned} \tag{1}$$

The general solution with all parameters specified as symbols can be derived using built in Mathematica command **Solve**.

```
s1 = Solve[equDC, \{Ib, Vce\}].
```

$$\left\{\left\{Ib \to \frac{-Rb1\ Vbe + Rb2\ Vcc}{Rb1\ (Rb2 + Re + b\ Re)},\right.\right.$$
$$\left.\left.Vce \to Vcc + \frac{(Re + b\ (Rc + Re))\ (Rb1\ Vbe - Rb2\ Vcc)}{Rb1\ (Rb2 + Re + b\ Re)}\right\}\right\}$$

Fig. 5. Solution of the set of the circuit equations of the Load-Line Analysis

Solutions are illustrated in Fig. 5.

Some parameters we can arbitrary chose, while other parameters we can retain in a symbolic form, **Rb2**, and use for the design or optimization. Suppose that the constrain is to have at collector the half of the supply voltage, that can be specified as unknown voltage **Vp**, equal to the voltage between collector and emitter for the solution **s1**.

```
Vp = Vce /. s1[[1]]
```

The new solution of this constrain is in terms of **Rb2**

```
s2 = Solve[Vp == Vcc/5, Rb2] // Simplify
```

Using chosen values of all parameters (for supply voltage = 15 V, collector resistor = 5 kΩ, emitter resistor = 120 Ω, current transistor gain = 100, the upper base resistor = 120 $k\Omega$), the voltage at the collector node can be presented as a function of **Rb2**, that is according to the solution **s2**.

```
{Rb2→(Rb1 (2 b Rc Vbe+Re Vcc+b Re Vcc))/(2 b Rc Vcc-Rb1 Vcc)}
```

The chosen values of all parameters are

```
subs1 = {Vcc→15, Rc→5000, Re→120, Vbe→8/10, b→100, Rb1→150000}
```

The numeric value of a resistor **Rb2** becomes 11.55 $k\Omega$. The standard value that we can accept is 11.55 $k\Omega$. The voltage at the collector node can be plotted as a function of the second base resistor **Rb2**, as it is illustrated in Fig. 6. Note that this is not a load-line characteristic, but the collector node is a rational function of **Rb2**, that is

```
(581800 - 35 Rb2)/(12120 + Rb2).
```

It is obviously from the Fig. 6 that the transistor is in active region for second base resistor **Rb2** values between 8 $k\Omega$ and 16 $k\Omega$. Notice that the classic design starts with assumption that transistor is in active region and at the end of the analysis is checking that the assumption is true.

5.2 Symbolic analysis using CAS

Similar approach is possible for transient simulation of common emitter amplifier from Fig. 1 and the frequency domain simulation of the circuit. However, from the educational viewpoint a

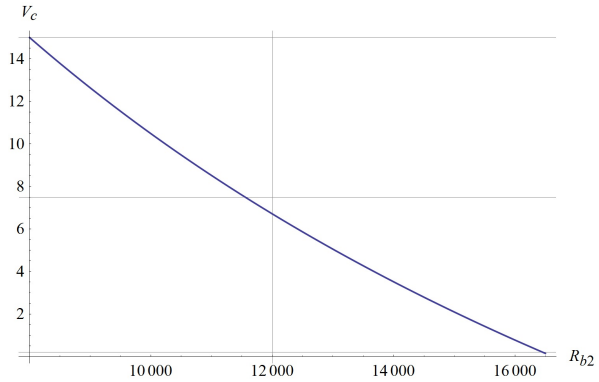

Fig. 6. The collector voltage as a function of Rb2

pictorial representation of the circuit is required. This can be done by entering the knowledge for drawing discrete electronic elements (**Get["drawfiltLC.m"]**) and known small-signal model of transistor amplifier, see Fig. 7. After reding drawing knowledge, the drawing grid should be defined using command **Table**

```
x = Table[i, {i, 400}]; y = Table[i, {i, 400}];
```

The schematic description is a list of elements (voltage source **VS**, current source **CS**, resistor **Res**, capacitor **Cap**, node **Node**, wire **Line**), where each element is also a list with all parameters that are required for drawing or solving linear circuit:

```
Shemantic {
  LCVS[x[[3]], y[[5]], y[[8]]," ", Vg, 1, ds/2,F],
  LCCS[x[[10]],y[[10]],y[[13]]," ",gm VBE,3,ds/2,F],
  LCRes[x[[3]], y[[8]], y[[11]]," ", Rg,1,ds/2,F],
  LCCap[x[[3]], y[[11]], y[[13]]," ",Cs,1,ds/2,F],
  LCNode[x[[7]], y[[13]], "BB", 1, 1,F],
  LCGrnd[x[[7]], y[[5]], 0, ds/2],
  LCLine[x[[3]], x[[13]], x[[7]], x[[13]]],
    ...
  LCNode[x[[10]], y[[10]], " ", 1, 1, F],
  {}
};
```

The schematic illustrated in Fig. 7 can be drawn using **Show** command.

```
Show[Schematic/. smenadrawLC /. {ds→5, F → 10}] /. subsRequ
```

Fig. 7 is illustrating the equivalent small-signal model of the amplifier with common emitter and the negative feedback using Re. For setting circuit equation, the modified nodal analysis is used that is very similar to the Kirchhoff s current law (identify all nodes, describe all currents entering nodes - algebraic sum of the currents entering a node in a circuit is equal to zero). This way, by adding parasitic capacitances we will not increase the number of equations.

In the presented example, we can identify only three nodes that are transistor ports (base, emitter, and collector). The set of equations circuit of small-signal model of the common emitter amplifier is as follows:

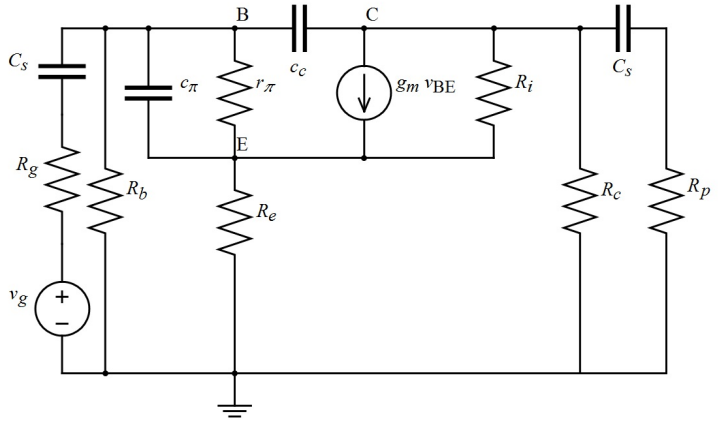

Fig. 7. Equivalent small-signal model of the amplifier with common emitter

```
equ = {(VB-Vg)/(Rg+1/(s Cs))+VB/Rb12+(VB-VE)/Rpi+
(VB-VE)/(1/(s Cpi))+(VB-VC)/(1/(s Cc))==0,
(VE-VB)/Rpi+(VE-VB)/(1/(s Cpi)) + VE/Rem-gm (VB-VE) +
(VE-VC)/Ri==0,
gm(VB-VE)+VC/(Rp+1/(s Cs))+VC/Rc+(VC-VE)/Ri +
(VC-VB)/(1/(s Cc))==0}
```

$$\begin{aligned}
&\left(Cs^2\, Rc\, Rp\, Rb\, s^2\, (Ri\, (cc\, r\pi\, s-\beta)+Re\, (1+r\pi\, c\pi\, s+cc\, s\, (Ri+r\pi+Ri\, r\pi\, c\pi\, s+Ri\, \beta)))\right)/\\
&\Big((1+Cs\, Rg\, s)\, ((1+Cs\, Rp\, s)\, (Ri\, r\pi+Re\, (Ri+r\pi+Ri\, r\pi\, c\pi\, s+Ri\, \beta))+\\
&\quad Rc\, (r\pi+r\pi\, s\, (Cs\, (Ri+Rp)+cc\, (Ri+Cs\, Ri\, Rp\, s))+\\
&\quad\quad Re\, (1+s\, (r\pi\, c\pi+cc\, (1+Cs\, Rp\, s)\, (Ri+r\pi+Ri\, r\pi\, c\pi\, s+Ri\, \beta)+Cs\\
&\quad\quad\quad (Ri+Rp+r\pi+Ri\, r\pi\, c\pi\, s+Rp\, r\pi\, c\pi\, s+Ri\, \beta)))))+\\
&\quad Rb\, \big((1+Cs\, Rp\, s)\, (Ri+Ri\, s\, (r\pi\, (cc+c\pi)+Cs\, (Rg+r\pi+Rg\, r\pi\, (cc+c\pi)\, s))+\\
&\quad\quad Re\, (1+r\pi\, c\pi\, s+cc\, s\, (Ri+r\pi+Ri\, r\pi\, c\pi\, s+Ri\, \beta)+Cs\, s\, (Ri+r\pi+Ri\, r\pi\, c\pi\\
&\quad\quad\quad s+Ri\, \beta+Rg\, (1+r\pi\, c\pi\, s+cc\, s\, (Ri+r\pi+Ri\, r\pi\, c\pi\, s+Ri\, \beta)))))+\\
&\quad Rc\, \big(1+s\, \big(r\pi\, c\pi+cc\, (r\pi+Ri\, (1+r\pi\, c\pi\, s+\beta))+Cs^2\, s\\
&\quad\quad (r\pi\, (Ri+Rp+cc\, Ri\, Rp\, s)+Re\, (Rp+r\pi+Rp\, r\pi\, (cc+c\pi))\, s+\\
&\quad\quad\quad Ri\, (1+cc\, Rp\, s)\, (1+r\pi\, c\pi\, s+\beta)+Rg\, (1+r\pi\, c\pi\, s+\\
&\quad\quad\quad\quad cc\, s\, (Ri+r\pi+Ri\, r\pi\, c\pi\, s+Ri\, \beta)))+Rg\, (Rp+Rp\, r\pi\, (cc+c\pi)\, s+\\
&\quad\quad\quad Ri\, (1+r\pi\, c\pi\, s+cc\, s\, (Rp+r\pi+Rp\, r\pi\, c\pi\, s+Rp\, \beta))))+\\
&\quad\quad Cs\, (Rg+Ri+Rp+r\pi+2\, Re\, (1+r\pi\, c\pi\, s+cc\, s\, (r\pi+Ri\, (1+r\pi\, c\pi\, s+\beta)))+\\
&\quad\quad\quad s\, ((Rg+Ri+Rp)\, r\pi\, c\pi+cc\, (Rp\, r\pi+Ri\, (Rp+2\, r\pi+Rp\, r\pi\, c\pi\, s+Rp\, \beta)+\\
&\quad\quad\quad\quad Rg\, (r\pi+Ri\, (1+r\pi\, c\pi\, s+\beta)))))\big)\big)\big)
\end{aligned}$$

Fig. 8. Transfer function of small-signal model of the common emitter amplifier

The transfer function can be derived after using the command for solving system of equations:

```
s1 = Solve[equ, {VE, VB, VC}]
```

The derived transfer function is presented in Fig. 8.

5.3 Optimization using CAS

The squared magnitude response of the transfer function at an specific frequency from the frequency range of interest can be derived using the following command (for known numeric values of all parameters but keeping emitter resistnce as a symbol)

```
Am = ((AvAll /. s -> -s) (AvAll)) /. s -> I 2  30000 /. subsN
```

The squared transfer function in terms of emitter resistance is presented in Fig. 8.

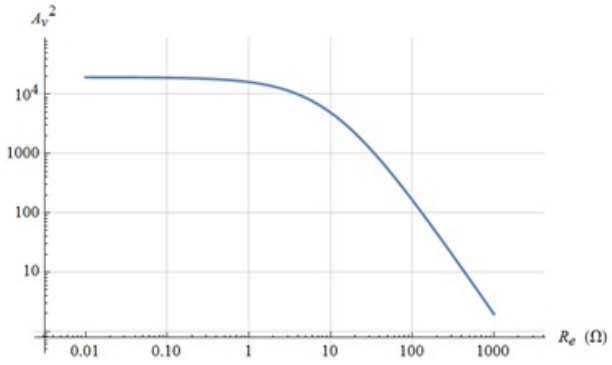

Fig. 9. Frequency domain simulation

Suppose that the requirement is to achieve the gain 10 of this amplifier (the squared gain magnitude is 100). From this constrain we can derive the value of emitter resistance.

```
solRe = Solve[Am == 100, Rem]
```

The solution is the positive resistance value that is Rem $\rightarrow$ 131. Substituting all values in transfer function, including **Rem**, the transfer function becomes a rational function in frequency. The plotted frequency response is presented in Fig. 10. This new approach presents that with computer algebra systems, such as Mathematica, it is possible to calculate in minutes or seconds the results that would be almost impossible by paper and pencil. Closed form expressions derived using this methodology can be used for optimization or an energy efficient design. In order to reduce the size of chips without needing significant design changes, the same methodology can be used for the design with programmable analog devices [5].

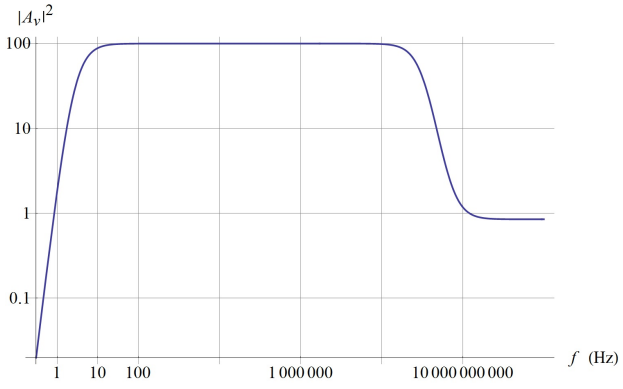

Fig. 10. Gain of the common emitter amplifier

5.4 Analysis of negative feedback circuits using CAS

When feedback is present the so-called closed-loop gain is sometimes not so simple to derive. There are some roles for efficient computation, but the derivation is usually depending on type of feedback and specific structure. One possible solution is to use CAS. The first step is to derive the input and output impedance of the circuit shown in Fig. 1 and the corresponding small signal model in Fig. 7. The input impedance is simply ratio of generator voltage and the current through the generator resistance. That follows from the same analysis in previous section. In the case of output impedance the generator should be placed at the output and new set of equations should be typed. The schematic is illustrated in Fig 11. Following the same procedure, the input and output impedances can be derived. Input impedance in terms of emitter resistance is presented in Fig. 12, while the output impedance in terms of emitter resistance is in Fig. 13.

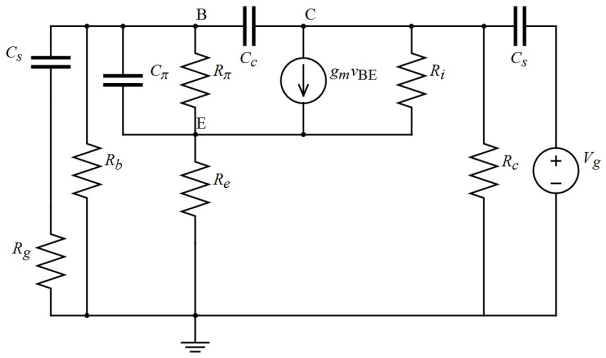

Fig. 11. Equivalent small-signal model for deriving output impedance

From figures we can conclude that input and output impedances are increasing functions of emitter resistance. That means, the type of feedback is series-series.

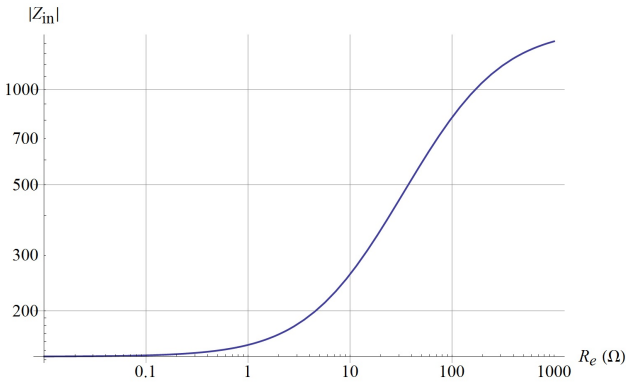

Fig. 12. Input impedance in terms of emitter resistance

Finally, the main result of this paper is that a straight-forward procedure can be used for typing equations by visual inspection of the schematic using basic electrical laws. The manual derivation is replaced using computer algebra system, and the results of analysis can be only interpreted from derived closed-form relations. By keeping values of some parameters in a

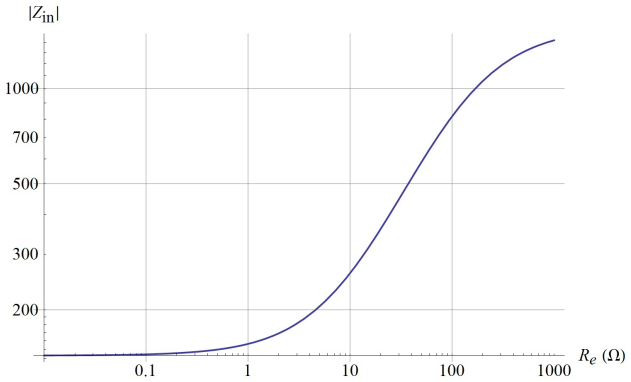

Fig. 13. Output impedance in terms of emitter resistance

symbolic form, optimization can be performed with additional constrains. This approach is usually called design-by-analysis because the starting point is known architecture of a system, and the values of the system parameters are results of symbolic or numeric optimization using CAS.

6 Conclusion

In recent years, classic electrical engineering education has focused largely on mathematics, manual derivations, and experiments using numeric software tools. The better insight in the design circuits and systems (understanding how electronic devices work, or how they can be used to create new more complex devices) is possible using computer algebra systems. The new approach is mainly targeted to combine understandings of simple devices, to integrate it into large systems, and to use computer algebra systems instead of manual derivation. Combining visual interpretation from mathematical models with simulating numeric tools, electronic kits, and the virtual instrumentation, the process of analysis and the design are integrated into unique development environment that can be used by students at the first and second year of electrical engineering, but also by researches without solid background in mathematics.

Acknowledgments. This work was supported by the Ministry of Education and Science of Serbia under Grant TR 32023.

References

1. Bert, D.: Raspberry Pi: Explore the RPi in 45 Electronics Projects. Elektor Publishing, Limbricht (2014)
2. Wolfram, S.: The Mathematica book, 5th Edition. Wolfram Media, Champaign, IL (2003)
3. Lutovac, D., Tosic, V. and Evans L.: Filter Design for Signal Processing Using MATLAB and Mathematica. Prentice Hall, Upper Saddle River (2001)
4. Hambley, R.: Electrical Engineering, Principles and Applications. Prentice Hall, Upper Saddle River (2011)
5. Lutovac, M., Pavlovic, V., Lutovac, M.: Automated Knowledge Based Filter Synthesis Using Gegenbauer Approximation and Optimization of Pole-Q Factors. Electronics and Electrical Engineering, 19 (2013)
6. Tosic, V.: Graph-theoretic formulation of equations for electrical circuits with Mathematica. The IPSI BgD Transactions on Internet Research, 6, 10-17 (2010)
7. Kraus G.: SPICE-Simulation using LTspice IV, `http://www.ieca-inc.com/images/Spice-Simulation_Using_LTspice_Part_1.pdf`, (2015)
8. Lutovac, D., Tosic, V.: SchematicSolver 2.3 for Mathematica 9. `http://www.wolfram.com/products/applications/schematicsolver/`, (2015)

9. Engelhardt M.: SPICE Differentiation. LT Journal of Analog Innovation, 25, 10-16 (2015)
10. Agarwal A., Lang J. H.: Foundations of Analog and Digital Electronic Circuits. Elsevier, Amsterdam (2005)
11. Froyd J E, Wankat P C, Smith K A. Five major shifts in 100 years of engineering education. Proceedings of the IEEE, 100, 1344-1360, (2012)

The Modeling and Control over Electrolytic Refining

Anna Sh. Lyubanova and Dmitrii N. Gron

Siberian Federal university,
Svobodny str. 79, 660041 Krasnoyarsk, Russia

Abstract. The paper discusses a mathematical model and algorithm of fuzzy control over the process of electrolytic copper refining. The mathematical model of control involves the regression model of the process, the limitations on the permissible values of technological parameters and the decision-making model. The algorithm of fuzzy control is developed on the base of the mathematical model and the Mamdani algorithm of fuzzy inference.

Keywords: mathematical model, decision theory, fuzzy control, copper electrolysis, process automation.

1 Introduction

The control is exerted over complex objects, as a rule, under a deficiency and inaccuracy of information, which can lead to wrong decisions and have grate consequences. In this situation, the current information technologies, in particular, decision support systems are best suited to the automated control of complex processes. The decision support systems should be used when the decision-making process cannot be completely formalized and implemented on a computer. The fundamental component of such systems is the mathematical model of the decision-making. Decision making is the process of selecting a possible course of action from a set of alternatives. For the control over technological processes, the set of alternatives is defined by the data on the technological parameters. However, there are many cases where the data may not be presented in a precise manner. An alternative way to deal with imprecise data is to employ the concept of fuzziness [5], whereby the vague data can be presented by fuzzy sets. The main advantages of methodologies based on fuzzy theory are that they do not require prior predictable regularities or posterior frequency distributions, and they can deal with imprecise input information containing feelings and emotions quantified based on the decision maker's subjective judgment. One of the most complicated processes is the production of the non-ferrous metals, particularly copper. In addition to the high complexity, the metallurgy technologies are characterized by the uncertainty of the data and the problems of gaining information. Therefore the concept of fuzziness and methodologies based on fuzzy theory are suited for the mathematical modeling of the control over such processes. In [2,3] a mathematical model of the process of electrolytic copper refining and a decision-making model are developed on the base of the experimental dependence of the output characteristics on the input data and control parameters [4]. The polynomial regression equations were constructed using the sample of the experimental data. This work proceeds the research started in [2]. The main goal of the work is to give a general concept (mathematical model) of the process of control over the electrolytic copper refining and create the algorhythm of fuzzy control based on this concept. The rest of the paper is organized as follows: In the following section, we discuss the mathematical model of the control over the electrolytic copper refining and the model of decision-making. On the base of these models we formulate the conception of the process control. In Section 3 we present an algorithm which realizes this conception using the methodology based on fuzzy inference.

2 The general conception and mathematical model of control

The process of electrolytic refining is carried out in electrolytic baths. The input parameters of the process are the concentration of the sulfuric acid in the electrolyte ($C_{H_2SO_4}$); the concentration of the copper sulphate in the electrolyte (C_{CuSO_4}); the copper concentration in the electrolyte (C_{Cu}); the current density (D); the rate of the electrolyte circulation (V). The output characteristics are the efficiency by current and the productivity of the electrolyzer.

The control process requires the special tools for defining the value of the control parameters. As such a tool the appropriate control algorithm can be created on the base of a mathematical model of electrolytic copper refining [3]:

$$\begin{aligned}\eta_{Cu} = {} & 885.52052 + 0.01869\ V + 0.01048\ D^2 - 5.79232\ D \\ & +0.000143\ C^2_{H_2SO_4} - 0.01231\ C_{H_2SO_4} + 0.000009\ C^2_{Cu} \\ & -0.00498\ C_{Cu} + 0.00035\ C^2_{CuSO_4} - 0.07688\ C_{CuSO_4},\end{aligned} \tag{1}$$

$$\begin{aligned}Pr_{Cu} = {} & -122.6664 + 0.0145\ V + 0.0024\ D^2 - 1.30096\ D \\ & +0.00036\ C^2_{H_2SO_4} - 0.09653\ C_{H_2SO_4} + 0.000044\ C^2_{Cu} \\ & -0.0027\ C_{Cu} + 0.00021\ C^2_{CuSO_4} - 0.0436\ C_{CuSO_4},\end{aligned} \tag{2}$$

$$D = IS^{-1}_{an} \tag{3}$$

where η_{Cu} is the unbiased estimate of mean of the current efficiency η (%); Pr_{Cu} is the unbiased estimate of mean of the productivity of the electrolyzer Pr (tonne per day); I is the current (kA); S_{an} is the anode space (m^2). The regression equations (1), (2) is constructed on the sample resulting from 67 measurements made in the electrolysis shop of "Open Joint Stock Company "Mining and Metallurgical Company "Norilsk Nickel". The range of the current efficiency was 84–92%. The general conception of the control over the electrolytic copper refining is as follows. The control is realized with respect to the controlling scheme in tune mode. The desired effect and the value of the output parameters are compared in the control facility. This allows determining the difference between the desired and real state of the object (given by the program). The control is based on the system of the residual functions $\eta_g - \eta$ and $Pr_g - Pr$ whereby the control parameters I and V are calculated:

$$I(\eta, Pr) = f_1(\eta_g - \eta, Pr_g - Pr), \tag{4}$$

$$V(\eta, Pr) = f_2(\eta_g - \eta, Pr_g - Pr). \tag{5}$$

Here, the subscript "g"means the given (preferred) value of the controlled parameter. The control problem is formulated as follows [3]: for given values of the parameters $C_{H_2SO_4}$, C_{CuSO_4}, C_{Cu} find the values of I and V which satisfy the inequalities

$$24 \le I \le 25, \tag{6}$$

$$12 \le V \le 30, \tag{7}$$

and minimize the residual functions:

$$\eta_g - \eta \to \min \qquad Pr_g - Pr \to \min\,.$$

In practice the last optimization criterion is usually replaced by the conditions when the value of the output parameters η and Pr lies in a sufficiently small neighborhood of η_g and Pr_g:

$$\eta_g - \eta \le \varepsilon_1 \qquad Pr_g - Pr \le \varepsilon_2 \tag{8}$$

where ε_1 and ε_2 - small quantities, $\varepsilon_1 \ge 0$, $\varepsilon_2 \ge 0$. In this case the control problem consists of finding the values of the control parameters I and V optimizing the residual functions $\eta_g - \eta$ and $Pr_g - Pr$ in sense of the criterion (8) under the constraints (6)–(7) provided that $C_{H_2SO_4}$, C_{CuSO_4}, C_{Cu} are given.

The electrolytic refining is a stochastic process and the equations (1) - (3) are a statistical model. Therefore when implementing the control over electrolytic refining in the tuning mode one needs take into account that the conditions (8) must be carried out in terms of the mean value: $M[\eta_g - \eta] \le \varepsilon_1$, $Pr_g - Pr \le \varepsilon_2$. It is clear that the exact verification of these conditions is impossible even with the best statistical estimates of the mean value. However, the statistical hypotheses for the unknown conditional mean can be tested for any fixed set of values of the input and control parameters.

$$H_0 : M[\eta_g - \eta] \le \varepsilon_1 \text{ and } Pr_g - Pr \le \varepsilon_2$$

$$H_1 : M[\eta_g - \eta] > \varepsilon_1 \text{ or } Pr_g - Pr > \varepsilon_2$$

The acceptance of the null hypothesis H_0 is enough to consider the value of the control parameters I and V as permissible.

In accordance with the principle of substantiation lack the estimates of the conditional means $M[\eta]$ and $M[Pr]$ and the coefficients in (1)–(2) obeys the normal distribution. By the maximum likelihood criterion [1], if the estimates of conditional means (1) and (2) satisfy the inequalities (8) independently of the conditional variance, then the probability of the untrue acceptance or rejection of the hypothesis H_0 is minimal. The control parameters should to be adjusted only when at least one of the conditions

$$\eta_{Cu} \ge \eta_g - \varepsilon_1, \qquad Pr_{Cu} \ge Pr_g - \varepsilon_2 \tag{9}$$

is violated. It is enough to choose values of the control parameters wherein the relationships (9) are fulfilled as equalities, that is D and V are the solutions of the system of equations

$$d_{Cu} + d_{11}D^2 + d_{12}D + d_{13}V = \eta_g - \varepsilon_1,$$

$$d_{Pr} + d_{21}D^2 + d_{22}D + d_{23}V = Pr_g - \varepsilon_2$$

whose left-hand sides are the expressions for η_{Cu}, Pr_{Cu} in (1), (2):

$$\begin{aligned} d_{Cu} = {} & 885.52052 + 0.000143\ C_{H_2SO_4}^2 - 0.01231\ C_{H_2SO_4} + 0.000009\ C_{Cu}^2 \\ & -0.00498\ C_{Cu} + 0.00035\ C_{CuSO_4}^2 - 0.07688\ C_{CuSO_4}. \\ d_{Pr} = {} & -122.6664 + 0.00036\ C_{H_2SO_4}^2 - 0.09653\ C_{H_2SO_4} + 0.000044\ C_{Cu}^2 \\ & -0.0027\ C_{Cu} + 0.00021\ C_{CuSO_4}^2 - 0.0436\ C_{CuSO_4}, \end{aligned}$$

$$d_{11} = 0.01048,\ d_{12} = -5.79232,\ d_{13} = 0.01869,\ d_{21} = 0.0024,$$
$$d_{22} = -1.30096,\ d_{23} = 0.0145.$$

Thus, solving this system of equations and taking into account (3) we can find the optimal values of I and V in sense of the criterion (8), that is, the optimal mode of operation for the electrolizer.

3 The mathematical model of decision-making

The mathematical model of the control over the electrolytic copper refining includes a mathematical model of decision-making along with the model of the process (1)–(3) and the technological constraints (6), (7). The model of decision-making involves the set of feasible alternatives Ω and a decision function. In our case the set of feasible alternatives consists of all admissible states S of the electrolyzer. Attributes of the state are the observable (input) and control parameters. According to the statement of the control problem the best alternatives are matched by the values of control parameters I and V fulfilling equations (9) for known observable parameters. Therefore the attributes I and V give the criteria of alternatives. Thus, the set of feasible alternatives Ω consists of all admissible pairs (I, V) satisfying the constraints (6) and (7). Let every pair (I, V) be denoted by the letter S again. To construct a decision function we define the Boolean function $f(S)$ so that $f(S) = 1$ if the predicate

$$P(S) = (\eta_{\mathrm{Cu}} \geq \eta_g - \varepsilon_1) \wedge (Pr_{\mathrm{Cu}} \geq Pr_g - \varepsilon_2)$$

takes the value "true"and $f(S) = 0$ when the value of P(S) is "false". The decision function is generated by the binary relation R_f according to the rule

$$S_1 R_f S_2 \Longleftrightarrow f(S_1) > f(S_2)$$

and takes the form

$$C(X) = \left\{ S \in X \middle| f(S) = 1 \right\} \tag{10}$$

on every subset $X \subset \Omega$. The set of non-dominated alternatives consists of all pairs $(I, V) \in C(\Omega)$. The condition $f(S) = 1$ is provided by the solution of the system (9).

4 The algorithm of the fuzzy control

The systems of fuzzy control or fuzzy inference are widely used for the control over technological processes and devices. The development of fuzzy inference systems involves a number of the stages realized by means of the fuzzy logic [5]. Information arrived at the system of the fuzzy inference involves the concentration of the copper in the anode and the electrolyte, the concentration of the sulfuric acid and the copper sulfate in the electrolyte. The current and the rate of the electrolyte circulation, being the control parameters, are generated at the output of the fuzzy inference system. The control conditions (9), the predicate $P(S)$ and the relation R_f generating the decision function (10) are meant in a fuzzy sense. The decision function is given in the form of the base of the fuzzy choice rules. For the fuzzy inference the Mamdani algorithm [5] including the following steps:

1) the formation of the base of rules for the fuzzy inference;
2) the fuzzification of the input variables;
3) the fuzzification of the output variables and the formation of subconclusions;
4) the aggregation of subconditions in the rules of fuzzy productions. The paired fuzzy logic operations are used for finding truth degree of the condition in every rule of fuzzy productions. The rule is considered active and used for further calculations if the truth degree of a condition is different from zero;
5) the activation of the subconclusion in fuzzy production rules is run by the method of min-activation; in this case active fuzzy rules of production are only taken into account for saving time;

6) the accumulation of the conclusions in the fuzzy production rules is achieved by the association of the fuzzy sets concerning with the subconclusions for the same output linguistic variables;
7) the defuzzification of output variable by the method of the gravity center.

The subconclusions in the fuzzy production rules are formed as follows. A triangular membership function is determined for each subinterval of the feasible range partition for the input and output parameters. The gravity centers are found for the fuzzy sets appearing in all subconditions of the rule Rk and in so doing the values of the input parameters are gotten. The obtained values of the input parameters are substituted into (9) in order to find the appropriate values of I and V. Next, the subdomains containing this values are defined. The value of every control parameter is the point belonging to the intersection of two subdomains. Matching the values of the membership functions of the two subdomains at this point one choose the subdomain whose membership function takes greater value. This subdomain is the support of the fuzzy set I_m or V_l defining the appropriate subconclusions of the rule R_k. The truth degree of conditions for every rule is calculated in the following way:

$$\mu_j = \mu^j_{\mathrm{Cu}} \cdot \mu^j_{\mathrm{CuSO_4}} \cdot \mu^j_{\mathrm{H_2SO_4}}$$

where μ_j is the truth degree of the j-th rule, μ^j_{Cu} is the truth degree of the predicate (C_{Cu} is C^j_{Cu}), $\mu^j_{\mathrm{CuSO_4}}$ is the truth degree of the predicate ($C_{\mathrm{CuSO_4}}$ is $C^j_{\mathrm{CuSO_4}}$) and $\mu^j_{\mathrm{H_2SO_4}}$ is the truth degree of the predicate ($C_{\mathrm{H_2SO_4}}$ is $C^j_{\mathrm{H_2SO_4}}$).

In fuzzy production rules the subconclusions are activated by the method of min-activation. The accumulation of conclusions of fussy production rules is resulted for each output variable as the association of fuzzy sets corresponding to this variable in all active rules.

The numerical experiment was performed with the data obtained on an industrial electrolyzer with 43 anodes and the total space 30 m^2. Indicators are taken with the following input parameters $C_{\mathrm{H_2SO_4}} = 100\mathrm{g/l}$, $C_{\mathrm{Cu}} = 80\mathrm{g/l}$, $C_{\mathrm{CuSO_4}} = 198\mathrm{g/m}$. The optimal values of the control parameters are $D = 272\mathrm{A/m}^2$, and $V = 19\mathrm{l/min}$. The average output current is increased from the average output current in real unit 87.46% to 88.99% and productivity - from 0,0498t. / Day. to 0,0506t. / day in one technological cycle. In accordance with the confidence interval [0.24, 0.34] for the discrepancy of the current efficiency one can assert that the absolute error of the current efficiency does not exceed 0.34% with probability 0.95. Therefore, the economic benefits for the current efficiency comprises 1.19% with the same confidence probability.

5 Conclusion

In this paper, we develop a general conception and the mathematical model of the control over the electrolytic copper refining. The algorithm of fuzzy control is developed on the base of this conception. Under the fuzzy control the range of fluctuation of the current efficiency and productivity decreases. Therefore the fuzzy control stabilizes the quality of control and allows to automatize the process of electrolytic refining and increase the rate of the current efficiency and productivity, i.e. to improve the quality of the electrolysis management.

References

1. 1. Voschinin, A.P., Sotirov, G.R.: Optimizatsiya v usloviyah neopredelennosti (Optimization under uncertainty). Publishing House of the MEI, Moskow. Technique. The People Republic of Bolgaria (1989)

2. 2. Gron, D.N., Gorensky, B.M.: Informatcionno-upravlyayushchaya sistema protcessom elektroliticheskogo rafinirovaniya medi (The information controlling system of the electrolytic copper refining). Journal of Siberian Federal University. The series "Technology". 4, 301-310 (2009)
3. 3. Gron, D.N., Lyubanova, A.S., Chentsov, S.V.: Povyshenie effektivnosti upravleniya protcessom elektroliticheskogo rafinirovaniya medi s pomoshch'yu SPPR (Improving the efficiency of the process control of the electrolytic refining of copper using decision support system). Fundamentalniye issledovaniya. 8 (Part 4), 822-827 (2013)
4. 4. Zotkov, O.M.: Elektroosazhdeniye medi: tekhniko-ekonomicheskii aspekt. Ch.1. Teoriya i praktika elektroosazhdeniya medi (Copper electrowinning: technical and economic aspect. Part 1. Theory and practice of copper electrowinning) The publushing company of Krasnoyarsk state University, Krasnoyarsk (1992)
5. 5. Zhang, H., Liu, D.: Fuzzy Modeling and Fuzzy Control. Birkhauser, Boston (2006)

About the Determination of Optimal Trajectory Condition of Stowage Material

S. Mustafin, Zh. Musina, A. Zeinullina

Institute of Information and Computational Technologies of MES RK,
Pushkin str.125, Almaty, Kazakhstan

e-mail:mustafinsal@mail.ru

Abstract. The method of forecasting of the filling concreting process according to the observational data of the stowing material state indices, carried out at earlier time points

Keywords: forecasting, process status, bookmark, safety, mining.

Introduction

The use of stowing on mining enterprises caused by the process safety of the performance of mining operations, preservation of buildings on the surface of the earth, safety and the environment control, etc. For this purpose, goaf is filled with stowing material, which is after reaching a certain state of the material should serve as supporting pillars. In mining stowing is defined as filling the goaf with stowing material, which is formed in the entrails of the earth as a result of mineral extraction. Stowing materials can be crushed rock formation, and production wastes. Stowing is solid if all the goaf is filled, and partial when certain parts (as tapes or layers) of it are filled. Depending on the way of transportation and stowage hydraulic, pneumatic, hydropneumatic, mechanical, self-flowing and manual are distinguished. The goals of the use of stowing material, depending on the purpose. Stowing used to control rock pressure, to reduce losses and dilution of extracted minerals in mining, to prevent mine fires, to reduce surface deformations of the earth and to protect the objects on the earth surface from damage, to improve the safety of mining operations, for improving the ventilation of underground workings, to reduce transport charges.

Formulation of the problem

Requirements for filling properties can be different and depend on its purpose. Thus, the requirements for filling used to prevent subsidence of the earth surface and thus the protection of buildings and structures is much higher and it is especially important to forecast its states of stowing, than in cases when e.g., stowing serves as the filler of voids and prevention of ore dilution and loss. Depending on the purpose and field development systems dry, hydraulic, hardening and other stowings are used. It is reasonable that properties and methods of their creation are different. At hardening stowing binder component is added, that significantly increases the cost of stowing due the high cost of binding material. This type of stowing greatly exceeds the cost of others and used in strictly defined cases and only under condition of full recoupment of materials and works on the stowing. There is a problem of definition of readiness of the state of hardening filling to perform intended functions [2]. Since the end of the 50s to solve the problem of the forecast of states one began to use different methods of modeling. According to these methods, complicated process is divided into physically different components, their separate study carried

out, and then their interference is determined by mathematical methods using computers. This is caused by the impossibility, in most cases, for reproducing in the lab at all the features the real process, accompanied by processes of mass and heat transfer. As is known, investigation of any process consists of sequential stages:

- □ statement of purpose and problem;
- □ study all the information about the process;
- □ construction of physical and chemical structure of the process model;
- □ preliminary construction of a mathematical model;
- □ decision of problems on optimal operating conditions of the process;
- □ carrying out computational experiments under different conditions;
- □ planning and implementation of natural laboratory experiment;
- □ comparison of the results of natural and computational experiments.

Naturally, the number of stages may be missing or combined into one. As a result of these stages there is a model of the process, ready for industrial implementation occurs. Therefore the development of process models becomes more actual. The models with the optimal design have particularly great importance, when it is required to estimate the optimal parameters of the behavior of processes under different conditions. In this case, the mathematical model has to satisfy a number of important requirements of adequacy, sufficient simplicity, stability relative to errors in the original data, simplicity, productivity, clarity, cost of obtaining initial data and a number of other conditions [5]. During the creation of hardening filling, which consists of filling and forming with the artificial array, due to shrinkage of stowing materials, is difficult to conduct field studies under industrial conditions on a wide range of indicators characterizing its condition. To assess the state of filling is possible to investigate its individual elements, drilling, for example, the stowing to obtain core samples at certain points in time (week, month, etc.) and, by examining their characteristics, make prediction of the entire filling state. Another way of predicting the state of filling can be physical modeling with equivalent materials. In this case, the object of the study can be examined by mechanical, ultrasound, electrical and other methods. Assessment of change in strength of ultrasound, electrical and thermal parameters of stowing with the lapse of time generally constitute predictive picture of filling. Attempts to solve the problem on the basis of similarity of macro kinetics theory and physical modeling were incomplete due to incompatibility of similarity conditions of chemical and physical components of the process. To solve the problems of macro kinetics should be known objective laws of proper chemical transformation, not distorted by the influence of transport processes, and laws of mass- and heat transfer. Laws of chemical conversion are expressed in the form of kinetic equations reflecting the dependence of reaction rate on the composition of the reaction composite, temperature, pressure, catalyst properties (for catalytic processes), and others. In all cases, forecasting assessment of filling seems necessary to measure at different time points stowing material parameters characterizing the static state and the dynamics [2].

Formal problem

The state of the stowing material will be understood as a set of values of the physical properties of the material at the moment t . These properties are mechanical, thermal, ultrasonic, electrical and other parameters of stowing material, measured at the moment t. Let at the input of the suggested assessment of the state of filling the values of input parameters X are fixed and the

output takes the values of output parameters Y. It is necessary to build a mapping of the $X \to Y$ so that correct output signal Y would be formed to each possible input signal of the X. The use of traditional methods of predicting the state of stowing by one feature has low reliability due to the complexity of the research object, which significantly reduces its scope in practice. In practice, the most common models are continuous regression ones, constructed for the entire set of objects. Their use is based on assumptions about the high-quality homogeneity of the considered set of raw data on the objects of research, the lack of breaks and jumps in a priori unknown dependence. The idea of the proposed method for forecasting the state of the stowing is the partitioning of the feature space, characterizing the state of stowing material at certain times, to classes of close objects in the feature space, for each class in these times its own prediction function is built. Construction of particular models on selected areas brings together particular forecasting models into an integrated model of predicting the state of stowing. In other words, an attempt to take into account the structural heterogeneity of the original data on the state of the object of study in time is engaged, i.e. approximate real dependence of status of stowing on the physical parameters of the array of piecewise linear function with respect to time [6]. A method that uses pattern recognition techniques can improve the accuracy of forecasts and get a model that more accurately expresses the predicted trajectory of changes in stowing material for a set of features of its state.

Conclusions

The formal statement problems of calculation of the state of filling during field development offered. A method for determining the state of filling, which allows identifying the readiness of stowing for further operation developed. Application of the developed methods to solve specific problems allows investigating in detail the state of filling with the development of its hardening, set the moment of its operational readiness.

References

1. Baikonurov O.A., Krupnik L., Melnikov V.: Underground mining with stowing. Alma-Ata,Nauka (1972)
2. Blekhman I.I., Myshkis A.D., Panovko Ya.G.: Mechanics and Applied Mathematics: Logic and features of the applications of mathematics. Moscow, Nauka (1990)
3. Zhuravlyov Yu.I., Zenkin A.I., Ryazanov V.V.: Prediction algorithm of states of manufacturing processes . Voprosy radioelektroniki. Moscow. 1, 32–34 (1976)

Integrated Approach for Implementing the Virtual Information Infrastructure of the Automated Process Control System

N.R. Musabekov [1], D.T. Kasymova [2], A.K. Muslimova [2], A.O. Utegenova [2], I.T. Utepbergenov [2]

[1] Kazakh National Technical University named after K. Satpayev, Almaty, Kazakhstan
[2]Institute of Information and Computational Technologies of MES RK,
Pushkin str.125, Almaty, Kazakhstan
nazarbek_2008@inbox.ru,dika.cat@mail.ru,almira_muslimova@mail.ru,utegenova77@mail.ru,irbulat@rambler.ru

Abstract. Virtualization has become firmly established in the modern practice of engineering training. We have developed a laboratory stand for discipline "Computer control of technical systems"and "Automation and control"for students of specialty "Space technology and technology"based on virtual instrument.

Keywords: virtualization, the information infrastructure, software simulator, the angular rate sensor, mathematical model, the lower level of APCS.

1 Introduction

The move to virtualized IT-infrastructure for enterprises can become an important step in the transition to the model of outsourcing and cloud computing. Indeed, ensuring easy moving of virtual machines within the corporate IT-infrastructure, you can take the next step by sending a spinning this virtual machine somewhere out in the clouds.

Modern standard automated technological process control system (APCS) contains 3 levels [1-3 and etc.]:

1. Lower level is a level of transducers of data point and actuated equipment;
2. Medium - level is a level of technological management subsystems, the central part of which is redundant, programmable logic controllers, for example, SIMATIC S7-400 (SIEMENS) and redundant stations, for example, ET200M (SIEMENS), that collect and process data from the lower level;
3. High level is a level of the operational control loop with remote control actuators, change of operation mode of the control subsystems, providing information on the monitors, audio and visual signaling, means of long-term and short-term data archive, and so on.

The analysis showed that nowadays the appliance of hosted options for infrastructure virtualization is gaining popularity. Despite of the fact that the market offers a wide range of cloud services for virtualization of APCS components, our search did not give in general a single integrated solution for APCS.

Therefore, the aim of the present work is to study the relevant problems, i.e. of virtualization of the APCS information infrastructure on the basis of an integrated solution.

In order to accomplish the above objectives require the following tasks:

1. Analysis of virtual infrastructure solutions for different levels of APCS;
2. Identifying the bottlenecks in implementation of cloud solutions of APCS;
3. Implementation of comprehensive solution of the virtualizing information infrastructure of APCS.

2 Analysis of hosted options for different levels of APCS

Infrastructure virtualization for the high level of APCS. For virtualization and following transfer to the high-level cloud of APCS there is suitable a sufficiently widespread cloud server, it is a reliable tool for transferring corporate infrastructure to the cloud [4-7]. It can be easily customized for a customer's requirements and is ready for use within a relatively short time. In addition the customer gets a guaranteed amount of RAM and processing power. Service rating is made for actual consumption, which frees the customer from overpayment for downtime. Thus, virtualization at this level of APCS optimizes the costs of information infrastructure. To work with virtual machines it is enough to provide access to the Internet from a PC, thin client, tablet or mobile phone. The equipment is located in a data center, which is equipped with all the necessary engineering systems and software to prevent equipment downtime. The clearly written system roles, as well as the physical location of the data on the server, allow avoiding of critical data's leakage to the outside environment.

Virtualization infrastructure for medium level of APCS. The foundation of the program-technical complex of the middle-tier of APCS is the redundant, programmable logic controllers and redundant stations. Virtual solution for microcontrollers are currently sufficiently developed and widely used in the APCS middle level, for example, virtual controller CoDeSys 3 [8,9]. Virtual controller not only allows you to debug the program, but also can be used as a target platform for interconnection of peripheral equipment and interaction with Com ports on PC.

Virtualization infrastructure for lower-level of APCS In fact, virtualization and cloud solutions of the lower level of APCS are based on creation of simulation software of the automated object and are mainly required for testing APCS projects and research purposes. To implement such decisions, some of the companies, manufacturers of PLC, provide the tools to simulate the input signals. For example, through programming environment Step7 of Siemens controllers there can be forcibly set any input in real PLC (to force). During using the Simulator controller input signal is simply set via the GUI. There is also software simulation of APCS for Factory Acceptance Test (FAT) simulator (SIM for factory acceptance) and they are classified according to the degree of accuracy of the simulation. Online there is quite a bit of information on the establishment of such simulators. Mostly used by Matlab + Simulink, LabVIEW, someone implements simulation inside the PLC. There are special development tools Win MOD, Mynah MiMiC, Siemens SIMIT, APROS.

There are two main options for lower-level implementation of APCS architecture: hardware and software - Simulator. Hardware virtualization.

Physical channels of the test system can be connected to the Simulator on an "input channel to the output channel"and vice versa [10] (Fig.1).

In this illustration there is used HMI package for SCADA systems for implementation of logic simulation of PLC and user interface. And so, a hardware simulator is suitable for testing small (in terms of the number of input/output signals) automation systems or systems consisting of small independent subsystems.

3 Development of a platform for lower-level virtualization of APCS

The necessity of the APCS lower-level virtualization is connected with need to automate the process of checking the management control system in the office, that is, in the absence of automation objects. Moreover, some situations may not be modelled on the real object for reasons of security, cost, etc.

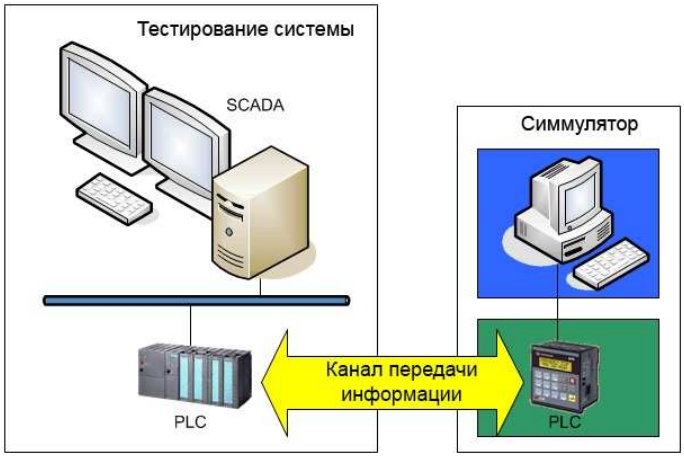

Fig. 1. Connection of the test system to the Simulator.

We have developed a laboratory virtual stand "Management of technical objects"for students of specialty "Space Hardware and Technology"and "Automation and control"on the base of the virtual appliance - Simulator for modeling and studying the complex object. Such objects are the sensors and actuators of lower level of APCS.

The Simulator is built on the same principle (Fig. 1) and can be used for debugging and testing of the developed software of APCS, as well as for the preparation of technological operators, who will work with her in the future.

For contacting the testable system the platform supports the DA OPC interface, which allows you to create configurations for two types of connections to hardware and software. In addition, support for connecting to data with SIMATIC PLCSIM allows easy debugging of Siemens S7 PLC software.

The advantage is that the software and hardware of test system, including channels for input/output signals shall be tested at the same time.

The GUI of the laboratory stand is a virtual device interface for the simulation and study of complex objects when performing laboratory work, such as the linear acceleration pick-up (accelerometers), a gyroscopic device for reproducing the object reference system (for example, using a gyro-stabilized platform) and allows to determine the angles of rotation and tilt of the object used to stabilize and control movement.

The graphical interface of the laboratory stand is a virtual instrument interface for modeling and studying complex objects when performing laboratory work, such as the linear acceleration sensors (accelerometers), gyroscopic device that plays on the object reference system (for example, using a gyro-stabilized platform) and allowing them to determine the angles of rotation and the inclination of the object used for stabilization and motion control.

Before students can be set a research task of varying difficulty, requiring for their solution wide range of statistical methods. The GUI is designed to simulate experiments and characterization of the object. In the main window of this application research topic is selected (by clicking on the corresponding button), further work is carried out in dialogue with the user, in which the simulator reflects on certain forms of "response"object to the user. The stand interface consists of the following steps:

1. Detection of the sensors input signals;
2. Assessment of the impact amplification of sensors output signals;

3. A block diagram of the sensor.

The user to conduct experiments on the stand (on the model) must learn the methods of research. The main menu the virtual stand consists of three laboratory works specifically implemented for systems' research and solving certain problems.

For the analysis of angular rate sensor (ARS) in mode dynamics in order to optimize its dynamic characteristics by clicking the mode "Block diagram of the sensor"the user is offered, the structure scheme. In the future the user will need to gather the scheme in the Matlab and to analyze the transition process.

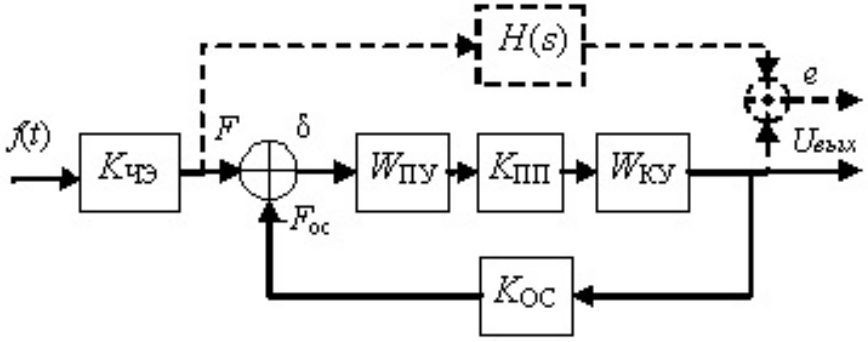

Fig. 2. Block diagram of the angular rate sensor

K_{cha} - transfer efficiency of sensitive device;

W_{pu} -transfer function of mobile node;

K_{pp} - transfer constant of capacitance displacement transducer;

W_{ku} - transfer function of the correction device;

K_{oc} - transfer constant of feedback link .

The overall transfer function, according to the organization chart is calculated as:

$$W(s) = K_{cha}\frac{W_{pu}K_{pp}W_{ku}}{1 + W_{pu}K_{pp}W_{ku}K_{oc}}$$

The basic mathematical models of sensitive device, transducer displacement, feedback main element are used in accordance with the theoretical terms of works [12].

The parameters of electrostatic vibration actuator, transmitter and transfer functions of the sensitive device and the mobile node are calculated on the basis of the system, similar to the microgyroscope LY530ALH firm Analog Device [13].

For real performance study of ARS and checking its specifications there is established an adequate model in visual environment Matlab Simulink when implementing previously calculated parameters.

A modeling diagram of angular velocity sensor in Matlab Simulink is represented on Fig. 3. The scheme provides a calculation of error in processing output step-wise signal by sensor structure. At that in the further software processing of errors' array received in the block e, an experimental value of the mean-squared error is calculated.

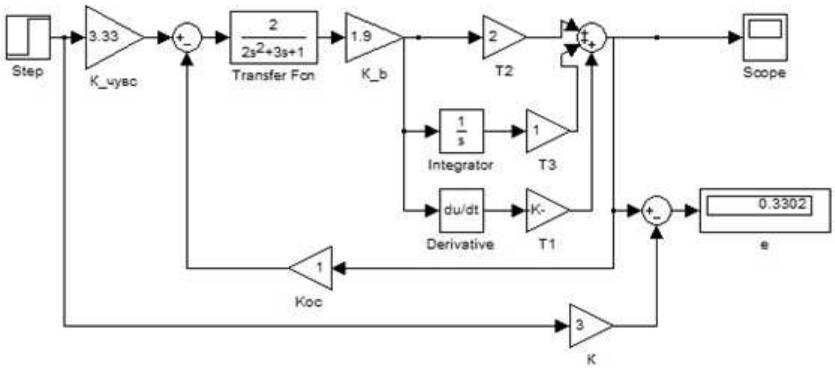

Fig. 3. Modeling diagram of sensor in Matlab Simulink

During the modeling there is received a timetable for the transition process at maximum input signal 150 degrees/sec (see Fig.4) showing aperiodicity of achievement of maximum output signal of 5 V at the time about $5.7 \cdot 10^7$ sec.

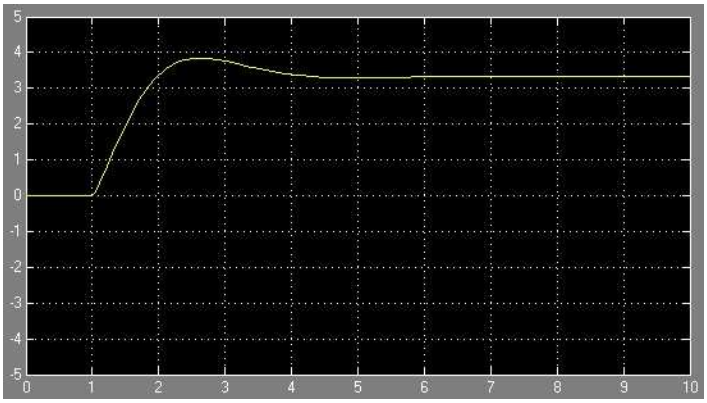

Fig. 4. Transition process of sensor

Thus, lower-level virtual object of APCS is received and ready for research.

4 Conclusion

1. There is established that there are technical facilities for integrated solution for virtualization of all three levels of APCS;
2. Virtualization of automated object on the lower level of APCS is necessary and appropriate for the test and research of APCS;
3. The Platform Simulator has been developed and researched to create virtual models of lower level of APCS on the example of angular rate sensor in dynamics mode;
4. Research results have been implemented in the educational process.

References

1. I.F. Borodin : Automation of technological processes. Moscow, (2007)

2. http://www.lanitnn.ru/page/406.html
3. B. Y. Sovetov Theoretical Fundamentals of automated control. Moscow, (2006)
4. http://www.intel.ru/content/www/ru/ru/cloud-computing/hybrid-cloud
5. http://selectel.ru/services/cloud/
6. A.G. Skhirtladze. Automation of production processes in engineering. Moscow: "Academy (2011)
7. I.D. Rudinskiy. Technology development of automated data processing systems and control Moscow: "Hotline Telecom"(2011)
8. http://plc24.ru/rabota-s-virtualnym-kontrollerom-codesys-3
9. M. Predko. PIC - Microcontrollers: Architecture and programming. Moscow: "DMK Press"(2009)
10. The simulator for testing SCADA software: Example http://habrahabr.ru/post/151125
11. Jeffrey Travis. LabVIEW for all // DMK Press. Moscow (2005).
12. V.K. Dedkov Prediction models of the individual reliability values. Moscow: Computer Center. A.A. Dorodnicyn RAS, (2003)
13. http:www.alldatasheet.com/datasheet346184/stmicroelectronics/ly530alh.htm

Discrete-Event Systems with State Observation Properties Studying

Nadezhda Nagul*

Matrosov Institute for System Dynamics and Control Theory,
Lermontov str., 134, 664033 Irkutsk, Russia
sapling@icc.ru
http://www.idstu.irk.ru

Abstract. The paper shows an application of the method of logical-algebraic equations (LAE-method) to the study of properties of controlled discrete-event systems (DESs). To apply the method we treat a property of the system under consideration as a property of an algebraic system. Then the algorithm is applied in order to construct the conditions of preservation properties of many-sorted algebraic systems in terms of morphisms. We address the issue of applying the LAE-method to DESs with partial observability in order to study their properties. Using the LAE-method, we examine some language properties preservation and demonstrate how the studying DES with state observation can easily be reduced to the same problem in the context of event observation.

Keywords: discrete-event systems, supervisory control, logic-algebraic equations, morphisms, properties preservation.

Introduction

This paper presents an application of the method of logical-algebraic equations (LAE-method) to the field of studying the properties of discrete-event systems (DESs). The LAE-method is the method of mathematical systems theory which lies at the intersection of system dynamics, algebra and logic and serves to synthesize criteria for preservation the properties of systems connected by special mappings — morphisms. One of the main applications of the preserving criteria thus obtained is the reduction of studying some complex system to studying a much simpler one.

To exploit the LAE-method we treat the property of the system under consideration as the property of an algebraic system. Due to the complex nature of dynamical systems the process of algebraizing of their models usually leads to many-sorted algebraic systems (MASs) where the basic sets have the meaning of a state space, a time scale, etc. Next we consider either of the logical-algebraic equations (LAE) $\mathcal{X}\&\mathcal{P} \to \mathcal{P}'$ and $\mathcal{X}\&\mathcal{P}' \to \mathcal{P}$ where $\mathcal{P}'$, $\mathcal{P}$ are known, and $\mathcal{P}$ is a formula predicate which describes the property of the system S, $\mathcal{P}'$ is a property of another system (S'), and $\mathcal{X}$ is the subject searching. The second equation corresponds to the preservation of properties in the direction which is opposite to mappings acting from S to S'.

Next we employ the algorithm of constructing the solution $\mathcal{R}$ of the chosen LAE which in place of $\mathcal{X}$ guarantees preserving the truth values of formula predicate under mappings of many-sorted algebraic systems to each other [7]. $\mathcal{R}$ is of the meaning of preservation operations and relations only, i.e. having the form of traditional morphisms. Morphisms of dynamical systems proved to be especially useful, for example, for those procedures of studying stability and other dynamic properties that require changing variables, since we should ensure that the property under consideration in old variables is equivalent to that one in new variables, or at least guarantee its unidirectional preservation.

* The reported study was partly funded by RFBR according to the research projects No. 14-07-00740-a, 14-07-31192-mol-a.

In this paper the LAE-method is applied to study the properties of DESs with partial observability. DESs model the system evolution by considering the occurrence of some event sequences. The development of DESs theory is driven by the rapid progress of manufacturing systems and communication networks, technological processes, transportation networks and others, primarily man-made systems. To regulate DES behavior the Ramadge-Wonham framework of supervisory control is commonly used, which adopt ideas from logic, language and automaton theory. In supervisory control theory the concept of observability plays an important role. Based on practical constraints, a subset of observable events is distinguished from all events. Observability was initially defined for event observation only, no state observation was explicitly considered. Using the LAE-method, we will demonstrate that the studying DES with state observation can be reduced to the same problem in the context of event observation.

DES with State Observation

In [2] the supervisory control problem of DES with state and event observation was introduced. We borrow that formulation with some changes and necessary corrections. Let

$$\mathcal{G} = (Q, \Sigma, \delta, q_0, Q_m) \tag{1}$$

be a discrete event system modeled as a generator of formal language [10]. Here Q is the set of states q; Σ the set of events; δ: $\Sigma \times Q \to Q$ the transition function; $q_0 \in Q$ the initial state; $Q_m \subset Q$ the set of marker states. As usual, Σ^* denote the set of all strings over Σ, including the empty string ε. In general, δ is only a partial function. Language generated by $\mathcal{G}$ is

$$L(\mathcal{G}) = \{w : w \in \Sigma^* \text{ and } \delta(w, q_0) \text{ is defined}\}$$

while language marked by $\mathcal{G}$ is

$$L_m(\mathcal{G}) = \{w : w \in L(\mathcal{G}) \text{ and } \delta(w, q_0) \in Q_m\}.$$

The Ramadge–Wonham supervisory control framework assumes the existence of a means of control $\mathcal{G}$ presented by a *supervisor*. Let Σ_c be a controllable event set, $\Sigma_{uc} = \Sigma \setminus \Sigma_c$, $\Sigma_c \cap \Sigma_{uc} = \varnothing$. The supervisor switches control patterns so that the supervised discrete event systems achieve a control objective described by some language K. Formally, a supervisor is a pair $\mathcal{J} = (\mathcal{S}, \phi)$ where $\mathcal{S} = (X, \Sigma, \xi, x_0, X_m)$ is a deterministic automaton with input alphabet Σ. S is considered to be driven externally by the stream of event symbols (words) generated by $\mathcal{G}$ (i.e. words from $L(\mathcal{G})$), while $\phi : X \to \Gamma$ is a (total) function that maps supervisor states x into control patterns $\gamma \in 2^{\Sigma}$. If S is in state x, the events $\sigma \in \Sigma_c$ of $\mathcal{G}$ are subject to control by $\phi(x)$. If $\sigma \in \phi(x)$, then σ is enabled, while if $\sigma \notin \phi(x)$ then σ is disabled (prohibited from occuring). Note that, unlike DES models with *forced events* [4], enabled events should not necessary occur. It is obvious that ϕ is the state feedback map. Because uncontrollable events cannot be disabled, it is required $\Sigma_{uc} \subseteq \gamma = \phi(x)$.

Let in addition Σ_o be an observable event set, $\Sigma_{uo} = \Sigma \setminus \Sigma_o$, $\Sigma_c \cap \Sigma_{uo} = \varnothing$ [5]. The supervisor observes only events from the observable event set and, basing on this information, controls events in a controllable event set, by disabling them. The observation function is defined as the projection $P : \Sigma^* \to \Sigma_o^*$. Since S is driven by the words of $L(\mathcal{G})$ along with restrictions on observation, a supervisor is characterized by a mapping $\gamma_e : P(L(\mathcal{G})) \to 2^{\Sigma_c}$. Here $\gamma_e(P(s))$ is interpreted as a set of events enabled by γ_e after observing $P(s)$, $s \in L(\mathcal{G})$.

It is assumed that in addition to event observation, supervisor also possesses some information on the states of $\mathcal{G}$. It observes $y = h(q)$, where $h : Q \to Y$ is an output mapping from the set

Q to the output set Y. To formally define a supervisor with event and state observation, the observation is extended in order to include the state observation. For a given string generated by $\mathcal{G}$, there is a unique sequence of states visited by the string. The supervisor observes the corresponding state output as well as the observable events. Such observations are characterized by the extended projection $P_s : L(\mathcal{G}) \to ((\Sigma_o \cup \{\varepsilon\}) \times Y)^*$ defined as

$$P_s(\varepsilon) = (\varepsilon, h(q_0)),$$

$$P_s(s\sigma) = P_s(s)(P(\sigma), h(\delta(s\sigma, q_0))).$$

Here, unlike [2], we do not admit multiple occurring of the artificial events corresponding to the output. Only one pair $(\sigma, y) \in \Sigma \times Y$ is added to the observation sequence, since the addition of more than one output is redundant.

We modify $\mathcal{G}$ by extending its states and events as follows:

$$\mathcal{G}_s = (Q \cup Q', \Sigma \cup Y, \delta_e, q_0', Q_m), \tag{2}$$

where Q' is the "double" of Q, that is, for each $q \in Q$, there is a double q' corresponding to q, and Q' is the set of all doubles; δ_s: $(\Sigma \cup Y) \times (Q \cup Q') \to Q \cup Q'$ is the extended transition function defined as follows:

$$\delta_s(\sigma, q) = \delta(\sigma, q)' \text{ if } \sigma \in \Sigma, q \in Q,$$

$$\delta_s(y, q) = q \text{ if } y = h(q), q \in Q,$$

$$\delta_s(y, q') = q \text{ if } y = h(q), q' \in Q',$$

and everything else is undefined. Thus, the supervisor is able to take the information about the output $y = h(q)$ into consideration since an output y is inserted before each occurrence of event. The artificial events in Y are considered to be uncontrollable.

Now we can define the language generated and marked by $\mathcal{G}_s$ via the language generated and marked by $\mathcal{G}$. Let us define $e : \Sigma^* \to (\Sigma \cup Y)^*$ as follows:

$$e(\varepsilon) = h(q_0),$$

$$e(s\sigma) = e(s)\sigma h(\delta(s\sigma, q_0)),$$

which is different from [2] but we believe is more correct. Again, we do not repeat insertion of the output, since it is redundant. The following lemma may be easily proved in the same way as in [2].

Lemma 1. *If $L(\mathcal{G})$ is not empty, then*

1. $L(\mathcal{G}_s) = e(L(\mathcal{G}))$
2. $T(L(\mathcal{G}_s)) = L(\mathcal{G})$
3. $L_m(\mathcal{G}_s) = e(L_m(\mathcal{G}))$

where $T : (\Sigma \cup Y)^* \to \Sigma^*$ is the projection.

To find an existence condition for a supervisor based on $\mathcal{G}_s$, denoted γ_s, the desired language K is extended to K_s as

$$K_s = T^{-1}(K) \cap L_m(\mathcal{G}).$$

Lemma 2. *[2] Assume that $K = \overline{K} \cap L_m(\mathcal{G})$. Then*

1. $\overline{K}_s = T^{-1}(\overline{K}) \cap L(\mathcal{G}_s)$,
2. $K_s = \overline{K}_s \cap L_m(\mathcal{G}_s)$,
3. $T(\overline{K}_s) = \overline{K}$.

In [2] an existence condition for γ_s is expressed in terms of controllability and observability of K_s.

The Method of Logical-Algebraic Equations

To exploit the LAE-method we treat the property of the system under consideration as a property of an algebraic system. Due to the complex nature of dynamical systems, the process of algebraizing of their models usually leads to many-sorted algebraic systems (MASs) where the basic sets have the meaning of a state space, a time scale, etc. Moreover, we introduced [7] the notion of a general many-sorted algebraic system of finite type (GMAS)

$$\mathfrak{A} = \langle A, \Omega_F, \Omega_P, \Omega_E \rangle, \tag{3}$$

where $A = \{A_\lambda | \, \lambda = \overline{1,k}\}$ is a family of basic sets, $\Omega_F = \{\mathbf{F}_\beta^{n_\beta} | \, \mathbf{F}_\beta^{n_\beta} : S_{1\beta}[A] \times S_{2\beta}[A] \times \ldots \times S_{n_\beta\beta}[A] \to S_{n_\beta+1,\beta}[A],\ \beta = \overline{1,k_F}\}$ a set of functions, $\Omega_P = \{\mathbf{P}_\gamma^{n_\gamma} | \, \mathbf{P}_\gamma^{n_\gamma} \subseteq T_{1\gamma}[A] \times \ldots \times T_{n_\gamma\gamma}[A],\ \gamma = \overline{1,k_P}\}$ a set of relations, $\Omega_E = \{\mathbf{E}_\delta | \, \mathbf{E}_\delta \in U_\delta[A],\ \delta = \overline{1,k_E}\}$ a set of distinguished elements. Here elements of the set $\Omega_F \cup \Omega_P \cup \Omega_E$ are defined on extended Bourbaki steps over the family A, which were extended with the operation of the sequence forming.

Consider the set of k symbols $\sigma_a = \{a_\lambda | \lambda = \overline{1,k}\}$ $(k > 0)$ and the signature $\sigma_0 = \{\sqcup \times \sqcup, \mathcal{P}, \mathcal{N}\}$.

Definition 1. *We define a step scheme over σ_a (or just scheme for brevity sake) as follows:*

1. *for all $\lambda = \overline{1,k}$ a_λ is a scheme;*
2. *if S_1, S_2 are schemes then $(S_1 \times S_2)$ is a scheme;*
3. *if S is a scheme then $\mathcal{P}(S)$ and $\mathcal{N}(S)$ are schemes;*
4. *expression is a scheme iff it follows from the rules 1 — 3.*

$St[\sigma_a]$ will denote *the set of all steps* S over σ_a. Given a family of sets $A = \{A_\lambda | \, \lambda = \overline{1,k}\}$, for any step scheme $S \in St[\sigma_a]$ by induction on the construction of the scheme S we can naturally and unambiguously define the set $S[A]$ — *a step over the family* A (built with the scheme S). In this case:

1. symbols $a_\lambda \in \sigma_a$ are interpreted as sets $A_\lambda \in A$ $(\lambda = \overline{1,k})$;
2. the binary symbol $\sqcup \times \sqcup \in \sigma_0$ is interpreted as the operation of constructing Cartesian product, i.e. for any schemes S_1, $S_2 \in St[\sigma_a]$ we assume $(S_1 \times S_2)[A] = S_1[A] \times S_2[A]$;
3. the function symbol $\mathcal{P} \in \sigma_0$ denotes Boolean constructing, i.e. for the scheme $\mathcal{P}(S) \in St[\sigma_a]$ we assume $\mathcal{P}(S)[A] = 2^{S[A]}$;
4. the symbol $\mathcal{N} \in \sigma_0$ denotes a sequences set constructing, i.e. for the scheme $\mathcal{N}(S) \in St[\sigma_a]$ we assume $\mathcal{N}(S)[A] = (S[A])^{I\!N}$.

Hence, the classical concepts of scheme and Bourbaki step [1] are extended by introducing to σ_0 the extra character $\mathcal{N}$ and the corresponding interpretation of schemes. The set of all such steps, generated as previously described, is called *generalized scale of steps* over A and denoted by $St[A]$. Due to the generalization of a scheme and N. Bourbaki step concepts, it is easy to represent various models of dynamic systems and, in particular, DES, in the form of GMAS.

We describe the system property under consideration with a formula predicate $\mathcal{F}(\overline{x}) \stackrel{df}{=} \mathcal{F}(x_1, \ldots, x_p)$ of signature σ of the chosen GMAS $\mathfrak{A}$, where x_μ is a free variable, $\mu = \overline{1,p}$, $p \geq 0$. Without loss of generality, the formula $\mathcal{F}$ is considered to be formed of *literals* (*concluding statements*, or *c-formulas*) $\mathcal{F}^\nu$, that is, atomic formulas $\mathcal{F}_+^\nu$ or their negations $\mathcal{F}_-^\nu$, with help of connectives &, $\vee$, and type quantifiers $\hat{\omega}_\alpha \stackrel{df}{=} \forall z_\alpha : Z_\alpha \stackrel{df}{=} \forall z_\alpha (Z_\alpha \to \sqcup)$ (universal type quantifier), $\check{\omega}_\alpha \stackrel{df}{=} \exists z_\alpha : Z_\alpha \stackrel{df}{=} \exists z_\alpha (Z_\alpha \& \sqcup)$ (existential type quantifier), $\alpha = \overline{1,n}$, $\nu = \overline{1,M}$. Such formulas, unlike positive formulas used, for example, in general algebraic system theory, will be referred to

as *generalized positive formulas*. The notion of "negation normal form" (ref. for example, [3]) is the analog of the notion introduced except that generalized positive formulas are built with type quantifiers instead of classic quantifiers. Here Z_α are so called *type conditions*.

Let us have a family of mappings

$$\varphi = \{\varphi_\lambda | \varphi_\lambda : A_\lambda \to A'_\lambda, \lambda = \overline{1,k}\} \tag{4}$$

which map the basic sets of GMAS $\mathfrak{A}$ to the basic sets $A' = \{A'_\lambda | \lambda = \overline{1,k}\}$ of the GMAS $\mathfrak{A}' = \langle A', \Omega'_F, \Omega'_P, \Omega'_E \rangle$, which is of the same type. "Same type" means that powers of sets A and A', Ω_F and Ω'_F and so on accordingly match and step $S'_{1\beta}[A'](S'_{2\beta}[A'], T'_\gamma[A'], U'_\delta[A']$ respectively) is formed from sets A'_λ with the same scheme as $S_{1\beta}[A]$ ($S_{2\beta}[A]$, $T_\gamma[A]$, $U_\delta[A]$ respectively) from sets A_λ. The aim is to find the way of synthesizing formula predicate $\mathcal{F}(\overline{x})$ preserving conditions in respect to mappings (4).

Let us add to the signature σ a set of symbols $\sigma' = \sigma'_a \cup \sigma'_P \cup \sigma'_F \cup \sigma'_E$ which doubles the signature σ, $\sigma' \cap \sigma = \varnothing$. We build the formula $\mathcal{F}' = \mathcal{F}(P'/P, F'/F, E'/E, x'_\mu/x_\mu, z'_\alpha/z_\alpha)$ of the signature σ', where expression X'/X means the substitution of all entries of the symbol X for the symbol X', i.e. each symbol $P \in \sigma_P$, $F \in \sigma_F$, $E \in \sigma_E$, ..., is replaced with a new corresponding symbol $P' \in \sigma'_P$, $F' \in \sigma'_F$, $E' \in \sigma'_E$, ..., $\mu = \overline{1,p}$, $\alpha = \overline{1,n}$. The formula $\mathcal{F}'(\varphi^{|x_1|}(x_1), \ldots, \varphi^{|x_p|}(x_p))$ thus obtained describes the property $\mathcal{F}(\overline{x})$ which is the same type to the property of the system $\mathfrak{A}'$. Symbols $\varphi^{|x_i|}(x_i)$ denote *the canonical expansion of mappings (CEM)* φ (4) on steps $S[A]$ $|x_i|$, to which variables x_i belong.

Next we construct logic-algebraic equations (LAEs)

$$\mathcal{X} \& \mathcal{F}(x_1, \ldots, x_p) \to \mathcal{F}'(\varphi^{|x_1|}(x_1), \ldots, \varphi^{|x_p|}(x_p)), \tag{5}$$

$$\mathcal{X} \& \mathcal{F}'(\varphi^{|x_1|}(x_1), \ldots, \varphi^{|x_p|}(x_p)) \to \mathcal{F}(x_1, \ldots, x_p). \tag{6}$$

Preserving conditions in respect to mappings (4) in a direction which is coincide with the direction of mappings between systems are to seek as a solution of LAE (5), while preserving conditions in direction, opposite to the direction of mappings between systems, are to seek as a solution of LAE (6). The detailed algorithm for finding non-trivial solutions of the equations (5) and (6) is presented in [7]. The main steps of the algorithm are the following:

Step 1 Separate the components of the signature σ in concluding statements: on the base of $\mathcal{F}(\overline{x})$ construct a formula $\Psi(\overline{x}) = \Psi(x_1, \ldots, x_p)$ which does not contain functional symbols in c-statements.

Step 2 Generate the preserving conditions in terms of CEMs. Denote them $\mathcal{R}_1$ for LAE (5) and $\mathcal{P}_1$ for LAE (6).

Step 3 Split the formulas obtained on the pervious step to morphism-like conditions. Construct the formula $\mathcal{R}$ as a solution of the LAE (5) or $\mathcal{P}$ as a solution of the LAE (6).

Conditions $\mathcal{R}_1$ and $\mathcal{P}_1$, obtained at step 2, are already sufficient to guarantee preservation of the property $\mathcal{F}$. Yet, as usual, they are too lengthy to deal with, therefore at step 3 they are split into conjunctions of simpler formulas which include in some sense "minimal" number of existential quantifiers and concluding statements. In [6] we introduced the notion of "standard splitting" which assumes the splitting $\mathcal{R}_1$ (or $\mathcal{P}_1$) according to the number of concluding statements, and then each part thus obtained are split according to existential quantifiers. Standard splitting allows one to generate instead of, for example, $\mathcal{R}_1$ a set of formulas collectionwise sufficient for $\mathcal{R}_1$ (or $\mathcal{P}_1$). Each element represents a condition of preserving the single symbol of the signature, functional or relation, in accordance with the traditional definitions of morphisms of algebraic

systems. Next, the conditions obtained are to be interpreted in the language of application domain. Standard splitting leads to obtaining the simplest structure of final formulas. This, in turn, allowed us to define the classes of properties which are preserved by the morphisms of the same class [6].

Language Properties Preservation

In [2] it was proved that since state observation provides more information about the system, the existence condition for a supervisor with event and state observation is weaker than the existence condition for a supervisor with event observation only. Moreover, they can be expressed in terms of controllability and observability of K_s. Using the LAE-method we will examine the connection between properties, basic for control problems, of the languages K and K_s.

Let $L \subset \Sigma^*$. The *closure* of L is the set of all strings that are prefixes of words of L, i.e. $\overline{L} = \{s | s \in \Sigma^* \text{ and } \exists t \in \Sigma^* : s \cdot t \in L\}$. A language L is closed if $L = \overline{L}$. If $\mathcal{G}$ is any generator then $L(\mathcal{G})$ is closed. Let K be a nonempty language that describes, for instance, the control objective for the system modeled by $\mathcal{G}$. Note that in this case for the existence of supervisor it is required K to be $L_m(\mathcal{G}) - closed$, i.e. $K = \overline{K} \cap L_m(\mathcal{G})$ [10].

Controllability

The definition of controllability plays a key role in characterizing those languages that can be generated by closed-loop structures "plant–supervisor".

Definition 2. *K is controllable (with respect to $L(\mathcal{G})$ and Σ_{uc}) if*

$$\overline{K}\,\Sigma_{uc} \cap L(\mathcal{G}) \subseteq \overline{K}.$$

We have that K, thinking of it as the admissible behavior of the system, is controllable if occurring of any uncontrolled event after prefix of the word from K leads to a word from K, i.e. still admissible. Controllability property may be expressed with the formula of the first-order predicate calculus language

$$\forall w \in \overline{K}\, \forall \sigma \in \Sigma_{uc}\, (w \cdot \sigma \in L(\mathcal{G}) \to w \cdot \sigma \in \overline{K}). \tag{7}$$

We transform of the formula (7) to a generalized positive form and in the notation of the language $L(\mathcal{G}_s)$. This leads to the formula

$$\mathcal{F}_1 = \forall w_s \in \overline{K}_s\, \forall \sigma_s \in \Sigma^s_{uc}\, (w_s \cdot \sigma_s \notin L(\mathcal{G}_s) \vee w_s \cdot \sigma_s \in \overline{K}_s). \tag{8}$$

While in general DES should be treated as a many-sorted algebraic system, taking into consideration the sets of events, states and so on, it is obvious that the controllability property may be dealt with as the property of the single-sorted algebraic system of finite type $\mathfrak{A} = \langle A, \Omega_F, \Omega_P, \Omega_E \rangle$, where $A = \{\Sigma \cup Y\}$, $\Omega_F = \{\cdot\}$, $\Omega_P = \{\overline{K}_s, \Sigma^s_{uc}, L(\mathcal{G}_s)\}$, $\Omega_E = \emptyset$. But nevertheless it remains "general". Indeed, since $\overline{K}_s, L(\mathcal{G}_s) \subseteq (\Sigma \cup Y)^{I\!N}$, the scheme $\mathcal{N}(a)$ corresponds to the predicates $\overline{K}_s$ and $L(\mathcal{G}_s)$. Therefore the additional operation of a sequence forming is used to build $\mathfrak{A}$.

Let $\mathfrak{A}' = \langle A', \Omega'_F, \Omega'_P, \Omega'_E \rangle$ be the GMAS of the same type as $\mathfrak{A}$ with $A' = \{\Sigma\}$, $\Omega'_F = \{\cdot\}$, $\Omega'_P = \{\overline{K}, L(\mathcal{G})\}$, $\Omega'_E = \emptyset$. The family of mappings (4) then consists of the single function

$$\varphi : \Sigma \cup Y \to \Sigma. \tag{9}$$

Note that according to step 1 of the algorithm, a new variable will be added to eliminate entries of the symbol $\cdot \in \sigma_F$ from the concluding statements of $\mathcal{F}_1$. To omit bulky formal manipulations of steps 2 and 3 we are going to use results obtained in [6]. For any formula $\mathcal{F}(\overline{x})$ let $ex(\mathcal{F})$ ($all(\mathcal{F})$, respectively) denote a set of variables of existential type quantifiers $\check{w}_\alpha$ of $\mathcal{F}$ (universal $\hat{w}_\alpha$, respectively). Let Q_{ex} (Q_{all}, respectively) denote a set of predicates which form type conditions of quantifiers for the variables from $ex(\mathcal{F})$ ($all(\mathcal{F})$, respectively). In case of (8) $Q_{ex} = \emptyset$, $Q_{all} = \{\overline{K}_s, \Sigma^s_{uc}\}$. A set of predicate symbols which form c-statements $\mathcal{F}^\nu_+$, we denote as $pos(\mathcal{F})$, while a set of predicate symbols which form c-statements $\mathcal{F}^\nu_-$ we denote as $neg(\mathcal{F})$. In (8) $pos(\mathcal{F}_1) = \{\overline{K}_s\}$, $neg(\mathcal{F}_1) = \{L(\mathcal{G}_s)\}$. Next we form the sets

$$\Omega_{P+} = Q_{ex} \cup pos(\mathcal{F}) = \{\overline{K}_s\},$$

$$\Omega_{P-} = Q_{all} \cup neg(\mathcal{F}) = \{\overline{K}_s, \Sigma^s_{uc}, L(\mathcal{G}_s)\},$$

$$\Omega_{P^+_-} = \Omega_{P+} \cap \Omega_{P-} = \{\overline{K}_s\}.$$

Note that $\Omega_{P+} \setminus \Omega_{P^+_-} = \emptyset$. This fact allows us to apply results from [6] as follows.

Let $\mathcal{I} \subseteq \{1, \ldots, k\}$ be a subset of the index set of the family (4). Let $\mathcal{I}_{S[A]}$ denote the index set of basic sets A_λ which are used to construct the step $S[A]$ with the scheme S.

Definition 3. *$\mathcal{MI}^-_\leftarrow$ class is the set of those general positive formulas $\mathcal{F}(x_1, \ldots, x_q)$ which satisfy the following conditions:*

1) predicates in type conditions of quantifiers do not contain functional symbols;

2) $\mathcal{I} = \bigcup_{S[A]} \mathcal{I}_{S[A]}$, where $S[A] = |v|$, v are all variables z_β or x_μ that enter the equalities $z_\alpha \equiv v$, which are included in c-statements $=_-$ or type conditions of the universal quantifiers $\hat{z}_\alpha$, and z_α is a defining variable;

3) $\Omega_{P+} \setminus \Omega_{P^+_-} = \varnothing$.

Recall that z_α is called a *defining variable*, if in $\mathcal{F}(\overline{x})$ the domain of the type quantifier $\tilde{z}_\alpha$ is larger than the domain of the type quantifier of the variable v. If $v = x_\mu$ than defining variable is z_α. The $t_1 \equiv t_2$ denote any of the equalities $t_1 = t_2$ or $t_2 = t_1$, while symbol $=_-$ (resp. $=_+$) denote entrance of the symbol $=$ in positive (resp. negative) literal (c-formula) of the initial formula $\mathcal{F}(\overline{x})$.

The definition below generalizes for GMAS the notion of a morphism of single-sorted algebraic systems and the notion of a Q-morphism of models by R. Lyndon [9].

Definition 4. *Let Θ be a set of relation symbols, $\Theta \subseteq \Omega_P$. The family of mappings (4) is said to be a $\mathcal{I}$-injective Θ-morphism GMAS $\mathfrak{A} = \langle A, \Omega_F, \Omega_P, \Omega_E\rangle$ to $\mathfrak{A}' = \langle A', \Omega'_F, \Omega'_P, \Omega'_E\rangle$, if*

1) $\langle\varphi\rangle^{S_{n_\beta}\beta+1,\beta[A]}(\mathbf{F}^{n_\beta}_\beta(z_1, \ldots, z_{n_\beta})) = (\mathbf{F}^{n_\beta}_\beta)'(\langle\varphi\rangle^{S_{1\beta}[A]}(z_1), \ldots, \langle\varphi\rangle^{S_{n_\beta}\beta[A]}(z_{n_\beta}))$, $z_1 \in S_{1\beta}[A]$, $\ldots$, $z_{n_\beta} \in S_{n_\beta\beta}[A]$, $\beta = \overline{1, k_F}$;

2) $\langle\varphi\rangle^{U_\delta[A]}(\mathbf{E}_\delta) = \mathbf{E}'_\delta$, $\delta = \overline{1, k_E}$;

and for all $\gamma = \overline{1, k_P}$

$$\langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}(\mathbf{P}^{n_\gamma}_\gamma) \subseteq (\mathbf{P}^{n_\gamma}_\gamma)'$$

if $\mathbf{P}^{n_\gamma}_\gamma$ correspond to a predicate symbol $P_\gamma \in \Theta$,

$$\langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}(\mathbf{P}^{n_\gamma}_\gamma) = (\mathbf{P}^{n_\gamma}_\gamma)',$$

if $\mathbf{P}^{n_\gamma}_\gamma$ correspond to a predicate symbol $P_\gamma \in \Omega_P \setminus \Theta$, and mappings φ_λ are injective for all $\lambda \in \mathcal{I}$.

Theorem 1 ([6]). *Let $\mathcal{F}(x_1, \ldots, x_q)$ be a formula of the class $\mathcal{MI}^-_{\leftarrow}$ with free variables x_1, …, x_q, $q \geqslant 0$. Then satisfiability of $\mathcal{F}'(f^{|x_1|}(x_1), \ldots, f^{|x_p|}(x_p))$ implies satisfiability of $\mathcal{F}$ under $\mathcal{I}$-injective $\Omega_{P-} \setminus \Omega_{P^+_-}$-morphism of GMAS $\mathfrak{A}$ to GMAS $\mathfrak{A}'$.*

We are going to use Theorem 1 to construct conditions for the controllability property preservation. The formula (8) obviously belongs to $\mathcal{MI}^-_{\leftarrow}$ class. It does not contain equalities so the set $\mathcal{I}_{S[A]}$ in this case is empty. We want to show that (9) may be defined to be a $\Omega_{P-} \setminus \Omega_{P^+_-}$-morphism. Let us start with condition 1) of Definition 4. The functional symbol "$\cdot$" is associated with triple occurring of the scheme $\mathcal{N}(a)$, and for all $s_s, t_s \in (\Sigma \cup Y)^{I\!N}$ an equality

$$\varphi|^N(s_s \cdot t_s) = \varphi|^N(s_s) \cdot \varphi|^N(t_s).$$

should be valid, since the operation of concatenation is the same in both $\mathfrak{A}$ and $\mathfrak{A}'$. It actually has a place if we define φ as follows:

$$\varphi(\sigma_s) = \begin{cases} \sigma, \text{ if } \sigma_s \in \Sigma; \\ \varepsilon, \text{ if } \sigma_s \in Y, \end{cases}$$

where ε is an empty string, or null event. This definition obviously correspond to the projection $T : (\Sigma \cup Y)^* \to \Sigma^*$.

Since $\Omega_E = \emptyset$ for (8), we skip the condition 2) and consider elements of the set $\Omega_{P-} \setminus \Omega_{P^+_-} = \{\Sigma^s_{uc}, L(\mathcal{G}_s)\}$. To satisfy Definition 4, it is necessary to guarantee for the predicate $L(\mathcal{G}_s)$

$$\langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}(L(\mathcal{G}_s)) \subseteq L(\mathcal{G}),$$

where $\langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}$ is CEM of φ over the step $T_{1\gamma}[A] \times \ldots \times T_{n_\gamma\gamma}[A]$, while $T_{1\gamma}$, …, $T_{n_\gamma\gamma}$ are the step schemes associated with the predicate $L(\mathcal{G}_s)$. Since in this case this is the scheme $\mathcal{N}(a)$, and $\mathcal{N}(S)[A] = (S[A])^{I\!N}$, then

$$\langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}(L(\mathcal{G}_s)) = \langle\varphi\rangle^{(A^{I\!N})}(L(\mathcal{G}_s)) = \varphi|^N(L(\mathcal{G}_s))$$

therefore $\varphi|^N(L(\mathcal{G}_s)) \subseteq L(\mathcal{G})$ should be valid. It is actually has a place because elimination of symbols of the set Y from the words of $L(\mathcal{G}_s)$ leads to the words from $L(\mathcal{G})$. For the predicate Σ^s_{uc} it is necessary to guarantee

$$\varphi(\Sigma^s_{uc}) \subseteq \Sigma_{uc},$$

where $\Sigma^s_{uc} = \Sigma_{uc} \cup Y$. If we suppose $\varepsilon \in \Sigma_{uc}$, this inclusion is satisfied.

It remains to prove that an equality $\varphi|^N(\overline{K}_s) = \overline{K}$ for the predicate from $\Omega_{P^+_-}$. It is easy to do using the same scheme as in the proof of Lemma 2 [2]. Now we have shown that (9) is a $\Omega_{P-} \setminus \Omega_{P^+_-}$-morphism and by Theorem 1 we formulate the following proposition, taking into account that $\mathcal{F}_1(\overline{x})$ do not contain free variables.

Proposition 1. *If K is controllable then K_s is controllable.*

Definition 5. *$\mathcal{MI}^-_{\rightarrow}$ is a class of general positive formulas $\mathcal{F}(x_1, \ldots, x_q)$ which satisfy the following conditions:*

1) predicates in type conditions of quantifiers do not contain functional symbols;

2) $\mathcal{I} = \bigcup_{S[A]} \mathcal{I}_{S[A]}$, where $S[A] = |v|$, and v are all variables z_β or x_μ, which enter to equalities $z_\alpha \equiv v$, which are included in c-statements $=_-$ or type conditions of the universal quantifiers $\hat{z}_\alpha$, and z_α is a defining variable;

3) $\Omega_{P+} \setminus \Omega_{P^+_-} = \varnothing$.

Since the formula (8) does not contain equalities, it belongs to $\mathcal{MI}^-_{\leftarrow}$ class as well as $\mathcal{MI}^-_{\rightarrow}$, with an empty set $\mathcal{I}_{S[A]}$.

Definition 6. *Let Θ be a set of relation symbols, $\Theta \subseteq \Omega_P$. The family of mappings (4) is said to be a mighty $\mathcal{I}$-injective Θ-morphism GMAS $\mathfrak{A} = \langle A, \Omega_F, \Omega_P, \Omega_E\rangle$ to $\mathfrak{A}' = \langle A', \Omega'_F, \Omega'_P, \Omega'_E\rangle$, if 1) and 2) of Definition 4 is satisfied and for all $\gamma = \overline{1, k_P}$*

$$(\mathbf{P}^{n_\gamma}_\gamma)' \subseteq \langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}(\mathbf{P}^{n_\gamma}_\gamma)$$

if $\mathbf{P}^{n_\gamma}_\gamma$ correspond to a predicate symbol $P_\gamma \in \Theta$,

$$\langle\varphi\rangle^{T_{1\gamma}[A]\times\ldots\times T_{n_\gamma\gamma}[A]}(\mathbf{P}^{n_\gamma}_\gamma) = (\mathbf{P}^{n_\gamma}_\gamma)'$$

if $\mathbf{P}^{n_\gamma}_\gamma$ correspond to a predicate symbol $P_\gamma \in \Omega_P \backslash \Theta$, and mappings φ_λ are injective for all $\lambda \in \mathcal{I}$.

Theorem 2 ([6]). *Formulas of the class $\mathcal{MI}^-_{\rightarrow}$ are preserved under mighty $\mathcal{I}$-injective $\Omega_{P-} \setminus \Omega_{P^+_-}$-morphism of GMAS $\mathfrak{A}$ to GMAS $\mathfrak{A}'$.*

In irder to prove that (9) is also a mighty $\Omega_{P-} \setminus \Omega_{P^+_-}$-morphism it is sufficient to show that $L(\mathcal{G}) \subseteq \varphi|^N(L(\mathcal{G}_s))$ and $\Sigma_{uc} \subseteq \varphi(\Sigma^s_{uc})$. Since these inclusions are obviously satisfied, with the help of Theorem 2 we obtain

Proposition 2. *K is controllable if and only if K_s is controllable.*

The proposition states that controllability is not affected by the state observation. This fact was previously demonstrated in [2] but its proof considerably differs from the one presented here.

Observability

Definition of observability means that if two strings look the same, then they must be consistent in the sense that no conflict of one event continuable after one string but not continuable after the other should occur [5].

Definition 7. *K is observable (with respect to $L(\mathcal{G})$ and P) if*

$$\forall s, t \in \Sigma^* \, (P(s) = P(t) \to consis(s, t)),$$

where $consis(s, t)$ is true if and only if

$$(\forall\sigma \in \Sigma)(s\sigma \in \overline{K} \,\&\, t\sigma \in L(\mathcal{G}) \,\&\, t \in \overline{K} \to t\sigma \in \overline{K}).$$

The observability property may be described by the generalized positive formula

$$\mathcal{F}_2 = \forall s, t \in \Sigma^* \, \forall\sigma \in \Sigma \, (P(s) \neq P(t) \vee s\sigma \notin \overline{K} \vee \\ \vee \, t\sigma \notin L(\mathcal{G}) \vee t \notin \overline{K} \vee t\sigma \in \overline{K}) \, . \quad (10)$$

We will consider the observability property of the language $\overline{K}_e$ therefore let $\mathfrak{B} = \langle B, \Omega_F, \Omega_P, \Omega_E\rangle$ be the single-sorted algebraic system of finite type where $B = \{\Sigma \cup Y\}$, $\Omega_F = \{\cdot, P\}$, $\Omega_P = \{(\Sigma \cup Y)^*, \Sigma \cup Y, \overline{K}_s, L(\mathcal{G}_s)\}$, $\Omega_E = \emptyset$. Note that "predicates" $\Sigma \cup Y$ and $(\Sigma \cup Y)^*$ here are artificial elements which introduced to embody the formula in the language of GMAS. Let $\mathfrak{B}' = \langle B', \Omega'_F, \Omega'_P, \Omega'_E\rangle$ be the GMAS of the same type as $\mathfrak{B}$ with $B' = \{\Sigma\}$,

$\Omega'_F = \{\cdot, P'\}$, $\Omega'_P = \{\Sigma^*, \Sigma, \overline{K}, L(\mathcal{G})\}$, $\Omega'_E = \emptyset$. $\mathfrak{B}$ is related to the observability property of the language K described with the (10). Here $P : \Sigma \cup Y \to (\Sigma_o \cup Y)^*$, $P' : \Sigma \to \Sigma_o^*$. Again, the family of mappings (4) includes the single function $\varphi : \Sigma \cup Y \to \Sigma$.

In case of (10) and $\mathfrak{B}$ $Q_{ex} = \emptyset$, $Q_{all} = \{(\Sigma \cup Y)^*, \Sigma \cup Y\}$, $pos(\mathcal{F}_2) = \{\overline{K}_s\}$, $neg(\mathcal{F}_2) = \{=, \overline{K}_s, L(\mathcal{G}_s)\}$,

$$\Omega_{P+} = Q_{ex} \cup pos(\mathcal{F}) = \{\overline{K}_s\},$$
$$\Omega_{P-} = Q_{all} \cup neg(\mathcal{F}) = \{(\Sigma \cup Y)^*, \Sigma \cup Y, =, \overline{K}_s, L(\mathcal{G}_s)\},$$
$$\Omega_{P^+_-} = \Omega_{P+} \cap \Omega_{P-} = \{\overline{K}_s\}.$$

Though the formula (10) does contain an equality, the set $\mathcal{I}_{S[A]}$ is still empty. As in the case of (8), $\mathcal{F}_2$ belongs to both $\mathcal{MI}^-_{\to}$ and $\mathcal{MI}^-_{\leftarrow}$ classes.

Note that in the case of (10), two new variables will be added to eliminate entries of the symbols $\cdot, P \in \sigma_F$ from the concluding statements of $\mathcal{F}_2$, according to step 1 of the algorithm. In company with the condition on the symbol $\cdot$, previously obtained, for the functional symbols P and P' the equality

$$\varphi|^N(P(t_s)) = P'(\varphi|^N(t_s)) \tag{11}$$

is valid, what may be easily checked. We skip the trivial equalities $\varphi((\Sigma \cup Y)^*) = \Sigma^*$, $\varphi(\Sigma \cup Y) = \Sigma$, and consider predicate "$=$". We now state that

$$[P(s_s) = P(t_s)] \Rightarrow [P'(\varphi|^N(s_s)) = P'(\varphi|^N(t_s))].$$

Indeed, as far as the output symbols $y \in Y$ are inserted regardless of observability of the symbols from Σ, this implication is true.

Thus, all conditions of Definition 4 are satisfied and (9) is a $\Omega_{P-} \setminus \Omega_{P^+_-}$-morphism. Therefore, according to Theorem 1, if K is observable then K_s is also observable. Let us show that in the opposite direction property is not preserved. According to Theorem 2 it should be valid

$$[P'(\varphi|^N(s_e)) = P'(\varphi|^N(t_e))] \Rightarrow [P(s_e) = P(t_e)].$$

Let us show this is not the case. Indeed, let we are given with two words $s_s = y_1 a y_2 b_1 y_3 c$ and $t_s = y_1 a y_4 b_2 y_3 c$ where b_1 and b_2 are unobservable. Then $\varphi(s_s) = a b_1 c$, $\varphi(t_s) = a b_2 c$ and $P'(\varphi(s_s)) = P'(a b_1 c) = ac$, $P'(\varphi(t_s)) = P'(a b_2 c) = ac$. But $P(s_e) = y_1 a y_2 y_3 c$ while $P(t_e) = y_1 a y_4 y_3 c$. So we formulate

Proposition 3. *If K is observable then K_s is observable while the opposite is not the case.*

Conclusion

In this paper there has been demonstrated an application of so-called LAE-method to generating the conditions of system property preservation relative to its mapping into another system. This method combines algebraic and logical approaches and may be of considerable interest to the audience. The theorems concerning the classes of the properties which are preserved under the similar type morphisms allowed one to easily obtain preservation criteria, basing just on the structure of the formula and omitting some numerous formal manipulations.

Although the presented results are quite simple, they illustrate the power of the LAE-method. Note that it allows us to generate the properties preservation conditions for different dynamical systems. The only stipulation of its applicability is the issue of algebraization of dynamical system models. For example, in [8] we investigated this way the DES-model of the public railway transportation network. The interpretation of the generated conditions allowed to obtain criteria of non-deterioration of the properties of the timetable when new railway paths are added to the network.

References

1. Bourbaki, N. Theory of Sets : Springer Science & Business Media (2004)
2. Changwei, C., Lin, F., Lin, Z.: Why event observation: observability revisited. Discrete Event Dynamic Systems. 149, 127–149 (1997)
3. Feferman, S.: Harmonious Logic: Craig's interpolation theorem and its descendants. Synthese. 3(164), 341–357 (2008)
4. Golaszewski, C.H., Ramadge, P.J.: Control of discrete event processes with forced events. Proc. 26th IEEE Conf. on Decision and Control (December 1987), 247—251 (1987)
5. Lin, F., Wonham, W.M.: On observability of discrete-event systems. Information Sciences. 44(3), 173–198 (1988)
6. Nagul, N.V.: Classes of properties preserved under morphisms of generalizations of many-sorted algebraic systems in studying dynamics. Trudy Instituta Matematiki i Mekhaniki UrO RAN. 2014. 1(20), 185–200 (2014)(in Russian)
7. Nagul, N.V.: The logic-algebraic equations method in system dynamics. St. Petersburg Math. J. 4(24), 645–662 (2013).
8. Nagul, N.V.: Timetable properties preserving in one model of public transportation network. The modern technologies. The system analysis. The modelling. 4(28), 150–159 (2010)(in Russian)
9. Lyndon, R.C.: Properties preserved under homomorphism. Pacific J. Math. 9, 143–154 (1959)
10. Ramadge, P.J., Wonham, W.M.: Supervisory control of class of discrete event processes. SIAM J. Control and Optimisation. 25(1), 206–230 (1987)

Simulation of the Control System Withdrawable Sensor Logging Stations *

K.A. Ozhikenov [1], R.M. Utebaev [1], R.S. Ismagulova [1], A.K. Ozhiken [2], G.D. Aitzhanova [2]

[1] K.I. Satpayev Kazakh National Technical University
Almaty, Republic of Kazakhstan
[2] Al-Farabi Kazakh National University, Almaty, Republic of Kazakhstan
{alfred.hofmann,ursula.barth,lncs}@springer.com

Abstract. This work addresses the is sue of geological exploration and search of rich ore deposits.One of the effective methods is a use of logging principle, which consists in:a well is drilled in 25-35cm diameter, then the well side face is studied, i.e. the ore percentage is determined on the well height.The logging unit specified in this work consists of a main stepper motor, reducer, winch and a shell with telescopic sensor.The sensor consists of ionizing radiation sources, proportional counter used for detector, preamplifier and a sensor movement stepper motor. A motor implements an important function, consisting in lowering and lifting of the logging instrument into the well.Therefore, a motor control function composition, the movement of which will reconcile with the movement of sensor movement motor.All software and control tools will be in the computer (laptop) with a digital converter. The task of creating a new logging unit, based on the use of massive material x-ray, as well as automation of all processes on this machine with the use of modern achievements of information technology and computer technology is relevant for the following features:creation of high technologies for solution of the tasks of exploration and prediction of mineral deposits within the priority of scientific and technological development of the mining and metallurgical industry of the Republic of Kazakhstan;the absence in Kazakhstan and the need to build a mobile logging unit with automated control and information processing system, based on the advanced method of well logging. The process of work covered a research on the development of a new method and technology of computer control and information processing of logging unit, an X-ray radiometrical logging unit mathematical model and control function was developed.The results of the study can later be used in creating automated mobile logging unit that can be used in mining metallurgical industry.

Keywords: well logging measurements, logging, logging unit, processing and conversion of information, software and hardware.

Introduction

To enhancegeo logical works, it is necessary to make a sharp breakthrough in the sphere of exploration, engineering and development of innovative prognostic and search technologies for identification of new promising deposits of various metals with mandatory certification by field studies. The number of the se developments include modern logging units that allow to preciselyassess the ore grade and presence of mineral deposits, to predict the further technology of extraction and processing [1, 2]. The purpose is to implement the methods of optimal design and logging unit operation control using new in format i on and information technologies. Wide development and spread of information technologies, as well as the iravailability has led to integration of many branches of science and technology with computer technologies [3, 4, 5].The use of information technology in geophysical researches of wells will let [6, 7]: optimize the logging process by creating a mobile logging unit with automated control and information processing

* The articleis prepared with in the frames of the project "A development and design of a driving actuator and a controlledparallel manipulator of x-ray radiometric logging station on the principle of feedback in accordance with the Order of the Chairman of the Committee of science of MES RoK No.17dated 08.04.2013.

system, based on the advanced method of well logging, reduce the cost of well exploration, data transmission by wireless networks will not require a foundering of additional expensive cables, recording and monitoring of the logging process in a well in real time, remote control of a shell with telescopic sensor.

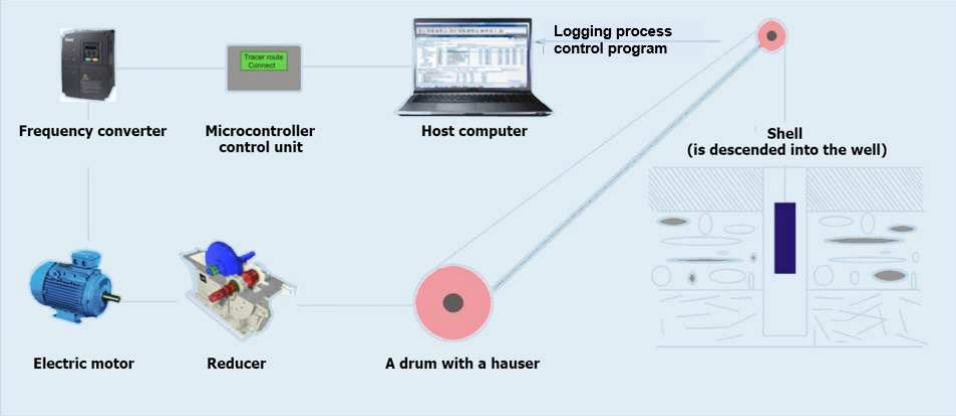

Fig. 1. Process functional diagram of XRLS

Logging unit designing goals. The initial goal in logging unit design is need to identify a set of functional and system requirements.On the part of a user the developed device is put the following requirements [8, 9, 10]: Indication of the current position of the probe (immersion depth); Alert on completion of measurements (beepnoise); An alert in detection of obstacles, closure of the flaps andlifting or stopping of the probe (a beep noise, engine stop); A developed logging system should have the following functions:

- Controlling the depth of device dismount into a well(lifting/lowering the probe);
- Control of the speed of cable movement, cable tension (measured parameters are displayed on digital light indicators);
- Sound light alarm, emergency stop (power off) when reaching the zero mark and/or programmable values in depth, as well as when reaching efforts on tension transducer of a predetermined value;
- A record on flash memory;
- Positioning of its location via GlobalPositioningSystem;
- Connection of personal computer (PC) with a measuring probe on a radio channel;
- Processing the flap position of the probe(open/close);
- Indication, charts;
- Rotation around the axis of the well, moving to required distance to the wall of the well.
- In accordance with the above functional and system requirements, x-ray radiometric logging station(XRLS)with computer control and data processing system was developed, which consists of three main parts [8, 9, 10]:power, electronic and software.The power partincludes a main motor, reducer, winch and a shell with telescopic sensor.A motor implements an important function, consisting in lowering and lifting of the logging instrument into the well.The sensor consists of ionizing radiation sources, proportional counter used for detector, preamplifier and a sensor movement stepper motor.Figure1showsa process functional diagram of XRLS.

Division of a logging unitintohardware (power, electronic) and software parts. This stage of workrequires a determination of the functions of hardware and software parts, so we allocate the functions. Functions of the hardware part:

- Display of the parameters on liquid crystal display;
- Control of the sound signal switching in overcharge;
- Control the opening/closing of the probe flap;
- Controlling the main engine;
- Positioning the sensor in the mine.

Functions of the software part:

- Receiving signals from the sensors;
- Beep;
- Determination of the depth;
- Detection of an obstacle;
- Determination of door position (open/closed).

At this stage of work it's necessary to make a general functional structure and algorithms of the software and hardware part operation.The end user must receive diagrame charts, displaying in formation on the content of minerals in drilling site, with subsequent comparison of the obtained results with the available spectrograms. This information is easy to process and present through a PC, and it can be receiveddescending a sensor(S) into a wellwith the help of a winch (W), to transmit information from the sensor to the PC data processing unit(DPU) is required. For automated reading of information, at different depths and in different directions a control unit(CU) is required, whichcontrols the positioning of the sensor in the well with a winch, and an angle of the sensor movement in a wellusing a stepper motor. Thus, the logging station can be divided into several functional blocks, shown in (figure 2). Let's consider each functional

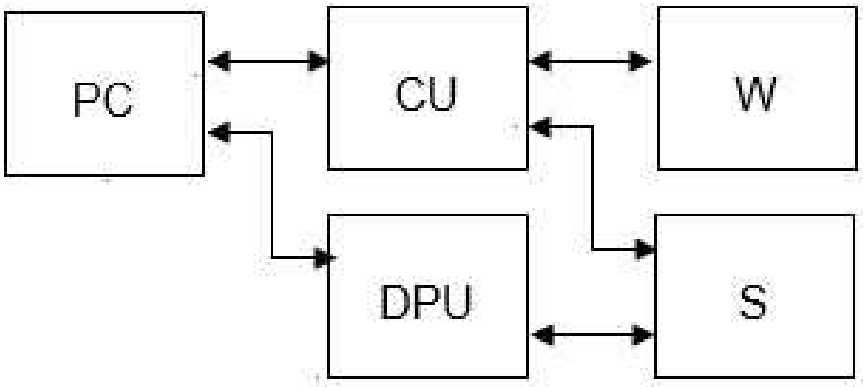

Fig. 2. Fumctional diagram

unit separately (figure 3). The whole station will consist of a stationary part and a movable module.Structurally terrestrial part consists of a PC, motor, micro controller (MC)ensuring a winch engine control;motor driver(MD), ensuring a connection of a motor with micro controller; a feedback(FB), ensuring a control of depth of immersion of the movable module, and a speed oflifting/lowering;radio transmitter for connection with a movable module. The movable module

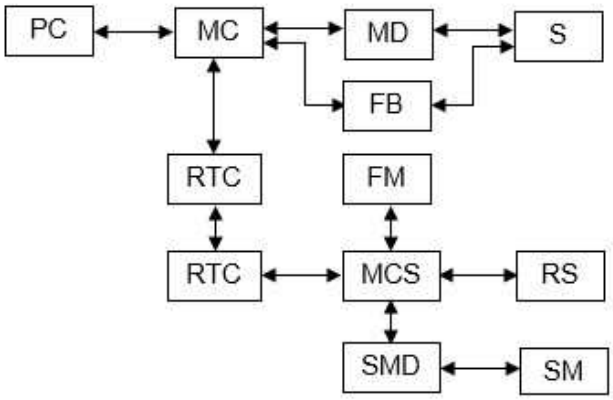

Fig. 3. A structure diagram of a logging unit

consists of a micro controller of the sensor (MCS), flash memory (FM), radioactive sensor (RS), stepper motor driver(SMD), stepper motor(SM), a radio transceiver (RTC). The micro controller reads the information from the sensor, records itinto flash memory, and transmits from RTC and MC to PC, controls the angle of sensor rotation by means of a stepping motor. X-ray

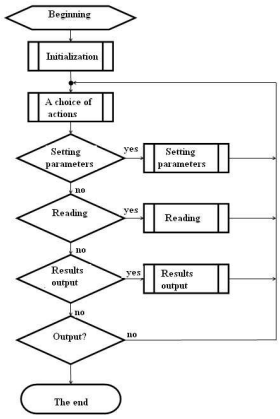

Fig. 4. A block diagram of the main algorithm

radiometric field instrument RPP-12is used as a sensor, which is designed for ore assaying in natural deposit(mine walls, open-cut mining benches, natural outcrops, etc.), in a loosened mined rock and coarsed-crushed sampling materials(hand specimen, core samples, etc.). The development of the software. For development of complex software it's necessary to develop an algorithm of its operation. The user will use the program on PC, which will represent an interface between the whole system and man. We divide this program into three large blocks:

– A setup of the whole system;
– Taking data from the well;
– Concluding results. A block diagram of the operation algorithm is presented in figure4.

During initialization, it's necessary to setinitial parameters of the program in the PC, to initialize the program on the micro controller, to establish a connection between MC and PC, check the system for performance. The block diagram of the initialization block is presented in figure5, a.

After initialization of the program the user will see the main interface of the program, where he can configure the system, run the program of wells research, or view the results of previous studies, from the file saved from the program, or saved to flash memory in a mobile module. The main parameters of the system are the measurement interval on depth, angle of rotation when scanning in different directions at the same depth, and initial and final depth of the scan. A block diagram of setting parameters is presented in figure 5, b.

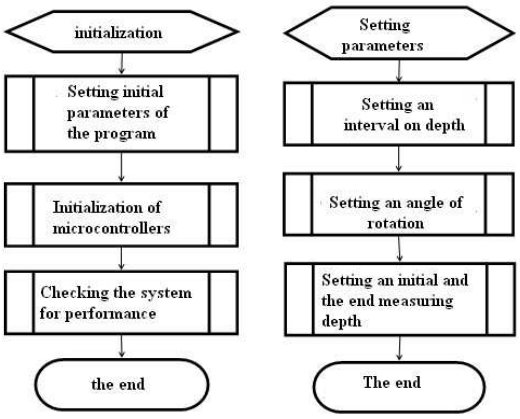

Fig. 5. A block diagram: a. -initialization of the program, b. -setting parameters

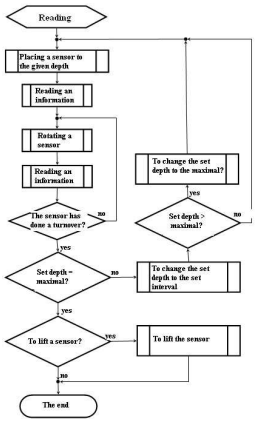

Fig. 6. A block diagram of readings

After starting the program of study, you cannot change the main parameters, you can only stop and set new parameters of research and continue from the set point. After starting the program the sensor is lowered to a predetermined starting depth, information is read from the sensor, is recorded on flash memory and transferred to PC. Then the sensor is rotated at a given angle and an information is read repeatedly, until the sensor returns to the start position. Then the sensor is lowered down to the specified interval, after which a reading and rotation procedure is repeated. It does so until the sensor is lowered to a depth set. Then the program stops and asks from the user the permission to lift the sensor, if the user refuses to lift the sensor, he should set parameters of a new study to implement additional study during the lifting of the sensor.

A block diagram of the readings is presented in figure6. While the information is read, data is placed in tables that are displayed in menu result output. Diagrams are built based on tables. This menu also includes results of previous studies, from the file saved from the program or saved to flash memory in a mobile module, and save the results to a file. A block diagram of result output is presented in figure 7.

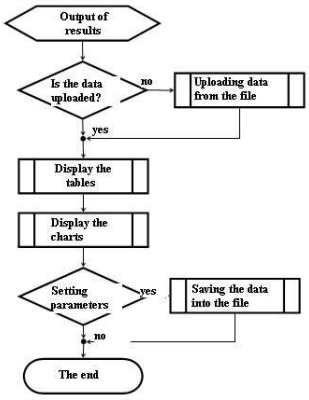

Fig. 7. A block diagram of result output

A technology of computer modeling of micro controller control system of a logging unit sensor actuator

This stage of work provides the main points that should be paid attention when design in gstepper motor(SM) control system based on PIC micro controllers [11, 12]. Let's consider a circuit diagram of micro controller control system of logging unit telescopic sensor actuator(figure8). The scheme is collected in intellectual environment of ISIS modeling of Proteus

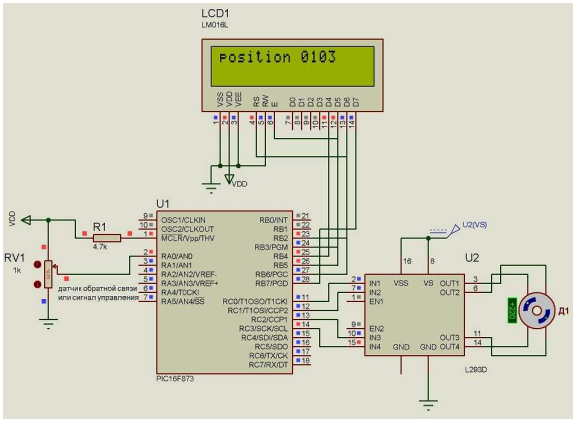

Fig. 8. A circuit diagram of stepper motor control system

package. The scheme consists of the micro controller U1, the driver SM U2, SM D1, LCD1 display and the variable resistor RV1 as a feedback or control signal. The applied micro controller PIC16F873 provides analog-to-digital conversion (ADC) analog input signal received at the port

AN0. The input signal can vary from 0to +5 B. After the ADC, the micro controller generates a control command in a sequence of connection of the windings SM D1 to the supply circuit and rotates the motor shaft at a given angle. Liquid crystal display (LCD) LCD1 displaying the processed control signal in decimal form indicates the position of the sensor. The drive control can be arranged in the form of a given program when upon receipt of actuator startup signal SM moves the sensor at a certain distance and after a specified time pushes it. In particular cases this may lead to failure of the mechanism, since without feedback the well wall roughness is not taken into account. To solve this problem it's necessary to used is tance measuring devices from a shell to a well wall at a given depth. An electric drive of logging unit telescopic sensor control unit is developed inProtonIDE. It is conditioned by the fact that Proton IDE is the simplest in solving such problems. It is a product of Mecanique company and designed to work in conjunction with micro controllers of Micro Chip company. After debugging a computer model of stepper drive we assemble real prototype of the electronic scheme(Figure9). Unlike computer models, the real

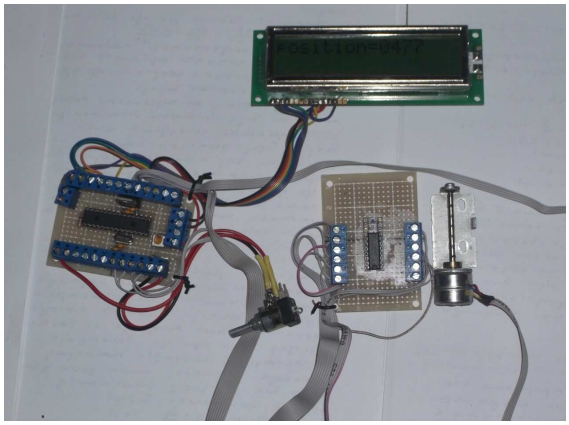

Fig. 9. Laboratory prototype of the logging unit electric drive

scheme include sadditional components, which are overdrawn in virtual environment in view of in sistence of their presence in the schemes. Thus, the real scheme uses a crystal resonatorto20 MHz, a voltage regulator for 5Vand terminals for node connections. External crystal resonator ensures a more precise operation of the micro controller. As there is one power source for stepper motor, display and micro controller, a voltage regulator provides a control scheme for stable 5 V power supply. Laboratory prototype uses miniature SM from computer floppy drive for flexible diskettes. It lets debug the control system on a real device, and at the same time not to spend a lot of financial resources for the experiment. The difference between the laboratory prototype and future trolley line is only in SM power, and respectively, in the amplifier. For miniatureSMmicrocircuitL293D driver is used, which provides bi-directional control of motors up to 4 watts. Main technical characteristics declared by the manufacturer of the microchip is presented in table 1. The result of modeling electromechanical processes in SM with regard to the technical characteristics presented in table 1, are shown in figure 10. In general, modeling results allow to evaluate the accuracy of the set movement, dynamic quality of the system, calculate the mechanical force in system element sand their impact on the main variables of the SM electromechanical processes. The conducted researches identified the possibility of creating a micro controller control system of electrical drive of logging unit telescopic sensor with the accuracy of a set movement. Table1 L293D technical characteristics

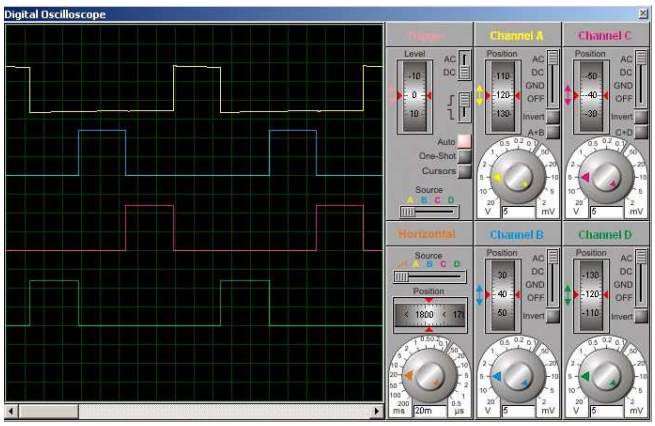

Fig. 10. Modeling results

Symbolic denomination	Parameter	Level
VSS	Logic Supply Voltage	36V
Vi	Input Voltage	7V
Ven	Enable Voltage	7V
Io	Peak Output Current(100msnonrepetitive)	1. 2A
Ptot	Total Power Dissipationat Tpins=90°C	4W
Tstg, Tj	Stora geandJunction Temperature	-40to150°C

The principle of building a connection between a shell with telescopic sensor and a host computer. In XRLSf or data transfer between the shell and host computerTCP/IP setup is used. In intranet, operating on the basis ofTCP/IP, the information is transmitted in the form of discrete units called IP packets or IP data grams [13, 14]. Due toTCP/IP software, host computer and logging shell, connected through a wireless Wi-Fi network are in the same intranet. Connections to intranet are identified by 32-bit IP addresses, which are expressed in the form of decimal numbers separated by dots [14, 15]. Figure 11 shows the scheme of connection to XRLS intranet blocks. Subscriber station 1 (Wi-Fiaccess point) is connected to hose computer via a twisted pair, then a signal through antenna is transmitted to a logging shell over the wireless network. Testing the micro controller control unit of shell with telescopic sensor in CAD. In addition to analysis of ore material in well, the development of information technology let use advanced visual observation transmitting the image in real time to the hose computer via a wireless Wi-Fi network. In this case, LiquidImageLIC727EGO camera, a pocket camera for extreme shootings, is used. A distinctive feature of this camera is the built-in Wi-Fi module, which allows to pre-set the angle and to select the most advantageous location. The camera allows to broadcast an image real-time to the host computer. Figure 12 shows the image transmitted by the camera to host computer.

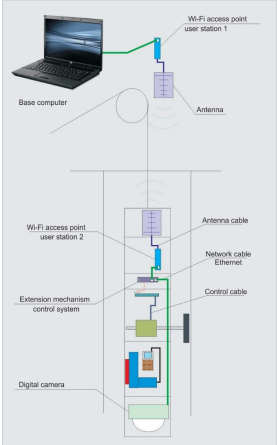

Fig. 11. A scheme of transferring data of hose computer andlogging shell In logging shell the signal is received by Subscriber station 2(Wi-Fiaccess point)and passes over twisted pair up to a digital camera and moveable mechanism control system.

Fig. 12. An image from the cameraat0. 27m, 24. 75mand76. 65m level, accordingly.

The camera also allowas to video record the well, it will help to clearly examine the entire column to determine the presence or absence of faults and fractures, as well as to know the water level, well depth and a clearance of drilled cement rock of themain water shut-off [14, 16]. The water level in the examined well is found at24.75m mark. At the level of 76.65m the ettles are seen on a well wall(figure 12).

Conclusions

- Mobile x-ray radiometric logging station that meets the functional system requirement of the end user is developed.
- A new method and technology of computer control and information processing of logging station are developed.
- A technology of computer modeling of micro controller control system of step perelectric drive of logging unit sensor is developed.
- The principle of construction of a wireless connection between a shell with with telescopic sensor and hose computer is developed.
- An algorithm and a control program of entire operation of logging unit electro mechanic system is developed.
- Results of the study revealed the possibility of creating a mobile logging station with micro controller control system of electrical drive of telescopic sensor of logging unit with the possibility of working out of a given displacement.

References

1. Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
2. May, P., Ehrlich, H.C., Steinke, T.: ZIB Structure Prediction Pipeline: Composing a Complex Biological Workflow through Web Services. In: Nagel, W.E., Walter, W.V., Lehner, W. (eds.) Euro-Par 2006. LNCS, vol. 4128, pp. 1148–1158. Springer, Heidelberg (2006)
3. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
4. Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: Grid Information Services for Distributed Resource Sharing. In: 10th IEEE International Symposium on High Performance Distributed Computing, pp. 181–184. IEEE Press, New York (2001)
5. Foster, I., Kesselman, C., Nick, J., Tuecke, S.: The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. Technical report, Global Grid Forum (2002)
6. National Center for Biotechnology Information, `http://www.ncbi.nlm.nih.gov`

Application of Programmable Logic Controllers for Efficient use of Photovoltaic Panels

Petrovic V. Vera, and Grujic Aleksandra

School of Electrical Engineering and Computer Science, Belgrade http://www.viser.edu.rs

Abstract. Great electricity consumption and population growth are forcing all countries to deal with the critical problem of reducing the stock of domestic fossil energy resources. Therefore, the proportion of electricity derived from renewable energy sources rapidly increasing in countries around the world. Serbia is yet to build significant capacity power plants that use renewable sources for electricity generation. One of the most promising forms of renewable energy is solar energy. Electricity generation, which is the most useful form of energy, is obtained by photovoltaic conversion of solar energy using photovoltaic cells that are grouped in photovoltaic panels. The highest utilization degree of solar energy is achieved when the angle of incidence of sunlight is 90°, which can only provide by mechanisms for tracking the position of the Sun because the Sun location constantly changing. This paper describes a mechanism that has the ability to monitor shifts the position of the Sun using the programmable logic controllers.

Keywords: photovoltaic panels, PLC, control, solar cells.

Introduction

The price of crude oil, which is constantly growing, increase in the price of natural gas for more than 50 percent and more expensive electricity is leading to increased use of renewable energy sources. To become independence on fossil fuel energy giants, many countries orient towards renewable energy sources, especially solar energy to. In this paper the conversion of solar energy into electricity using the photovoltaic effect, which is achieved through photovoltaic cells is presented. During lighting photovoltaic quantum sunlight induces electro-motor force to the ends of the photovoltaic cells.

There are many advantages of photovoltaic conversion. Solar cells directly convert solar energy into electricity without using mechanical parts, do not pollute environment, they need a minimum of maintenance by the service life of about twenty years. Their imperfection is current production of electricity, that's only in the period of radiation proportionally to the strength of solar radiation, density of power obtained is small (no more than about 100 W/m2), compared with the production of electricity from conventional sources their price is higher. It is the simplest and most attractive way of application of solar energy.

Direct conversion of solar energy into electricity is still too expensive to be economically competitive and therefore developed country governments subsidize the construction and use of power plants to generate electricity from solar energy. In this way the new materials are to be fined, and production technology become more simple. The current production technology of photovoltaic cells from silicon single crystals should become cheaper and increase the efficiency of these cells. The research of new materials, photovoltaic cells from polycrystalline and amorphous silicon, CdTe, GaAs and other semiconductor materials production and use concentrator photovoltaic systems, is likely to reduce the cost of solar electricity and enable its large-scale use.

Principle of solar cells

With the photovoltaic effect solar energy can be converted directly into electricity solar cells. When the solar cell absorbs solar radiation at its ends to a photovoltaic effect produces electromotive force and so the solar cells becomes a source of electricity. In Figure 1 is shown a cross-sectional Si solar cells.

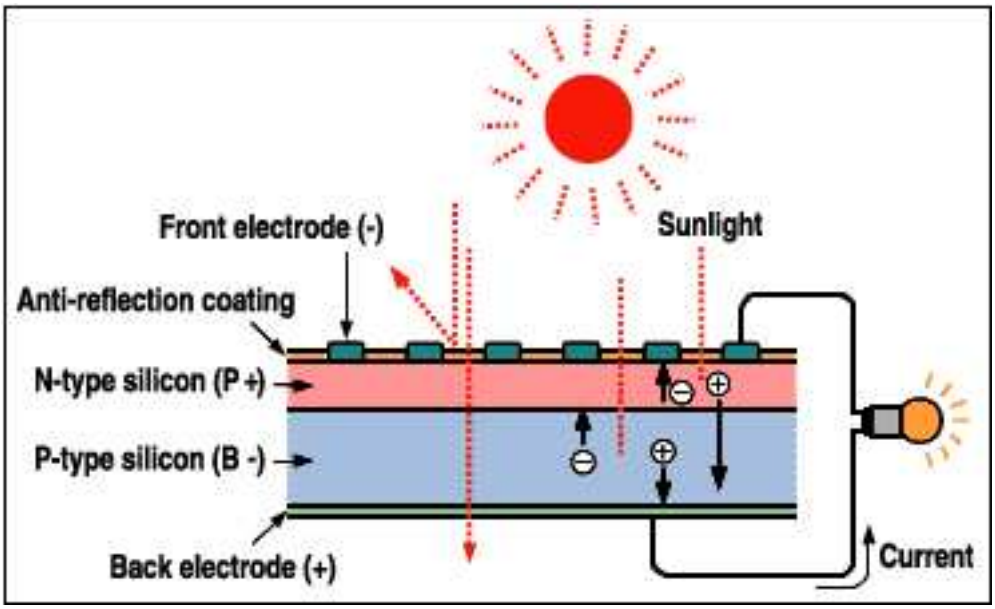

Fig. 1. Cross-sectional Si solar cells

To collect a charge that arose by absorbing photons from solar radiation, the front surface of the metal grille and rear side is covered with a metal contact. Lattice contact was made so as not to cover more than 5% of the area and he almost does not influence the absorption of solar radiation. The front surface of the cell can be covered with transparent and anti-glare coating that reduces the reflection of sunlight and thus increases the efficiency of the cell.

Solar cells produce a voltage of around 0.5 V with a density of 20 mA/cm. In order to obtain the appropriate voltage or power, the cells can be connected together in series and in parallel. This yields to a solar cell modules that are fixed and protected from weather and other influences. Modules can agree to each other in flat photovoltaic panels and panels together with other necessary elements (inverters, controllers, batteries, etc.) represents a photovoltaic system.

Solar cells should be prepared when illuminated, to take on its end a voltage. When illuminated absorbed photons produce electron-hole pairs. If absorption occurs the internal electric field, which exists in the impoverished area, separating the resulting electrons and holes (electron moving toward the p-side and cavity to the n-side). When each cell is illuminated on the p-contact part becomes positive, and on the n-part of the negative. When the solar cells connected to the external consumers, because PV cells will arise photocurrent, so the current flowing through the charge equal to the difference power LEDs and a photocurrent.

Materials for production of solar cells

To create mono crystals solar cells are used in silicon (Si), germanium (Ge), indium phosphide (InP), gallium arsenide (GaAs), cadmium sulfide (CdS), cadmium telluride (CdTe), aluminum antimonide (AlSb), gallium phosphide (GaP), cadmium selenide (CdSe) and others. Hence, it is used a pn-junction on the basis of one or more semiconductor materials from which to form the heterogeneous compounds, for example, a compound of Cu2S/CdS. However, today the most common solar cell based products, create single, polycrystalline and amorphous silicon, gallium arsenide (GaAs) and copper-sulfide/cadmium sulfide (Cu2S/CdS)

- Spectral sensitivity of solar cells

The solar cells are not equally sensitive to all wavelengths of the solar spectrum. The sensitivity depends on the nature of the conductor, present ingredients, technologies, etc. forming cells. Photons with short wavelengths are absorbed at the entrance of the solar cell, away from the p-n junction, and significantly affect the supply of solar cells. Photons with long wavelengths (infrared area) passing through the p-n junction, absorb at the bottom of the solar cell or reflected from the back electrode and do not contribute to significantly increase the current solar cells. Only photons with wavelengths in the area of maximum sensitivity of solar cells, significantly contribute to current solar cells.

- factors for efficiency of solar cells

On the efficiency of solar cells depends on several factors such as: a reflection on the surface of the solar cell, the losses by infrared, ultraviolet losses, losses due to the thickness of the solar cell, the losses due to factors voltages, losses due to recombination losses on the serial resistance.Optical reflection on solar cells depends on Micro roughness its surface. With increasing Micro roughness front surface of the solar cell comes to a reduction reflection. U to reduce reflections on the solar cell are applied appropriately anti-reflection coating. For solar cells with anti-reflection layers of optical reflection can be reduced to 3When mono crystalline (Si) solar cells, photons with energies above 1.1 eV generated photocurrent and surrender the excess energy that is pushing a single crystal is heated. In this way, about 33% efficiency solar cells is lost. When the absorption of solar radiation, electrons dont receive the entire amount of the absorbed energy in the material of solar cells. As a result, the voltage at the ends of the solar cell is less than expected. In this way, lost about 17% of the efficiency of solar cells. The generated electrons and holes in the solar cell when the solar radiation absorption have a certain lifetime after which can recombine, resulting in a 4% loss in its effectiveness. In the series resistance of the solar cell as a diode loses 1% of its efficiency. Some of these risk factors, loss of efficiency solar cells are determined by fundamental physical laws, so that they can be reduced. Losses of efficiency, which depend on the technology forming solar cells can be reduced. If the losses which depend on the technology reduce to the minimum, the maximum theoretical efficiency of crystalline Si solar cells amounted to 22%.

Photovoltaic panels

Photovoltaic source in the schemes shows the symbol for a p-n diode, but in fact it is the optical-mechanical-electrical structure, which consists of several parts. Electrical part of a photovoltaic panel includes all semiconductor elements on the board panel formed with contacts and cables. The power produced by a photovoltaic cell is small, realistic in practice more cells linked to the group, thereby forming a photovoltaic module. The modules are then combined bond and

build a photovoltaic panel that produces electricity, voltage and power of much greater intensity. The maximum output voltage of individual photovoltaic cells is about 600 mV, and the cells are serial link in order to obtain a desired voltage. The most commonly about 36 cells serially connects creating modules with nominal voltage of 12 V. Photovoltaic panel consisting of solar cells electrically connected in parallel combination of queues. Number of parallel connected cells makes sub-module, while the number of connected cells regularly makes serial string. The total amount of voltage at the output voltage is proportional to the number of series-connected cells, while the output current proportional to the number of parallel-connected cells.

Programmable logic controllers

Programmable logic controllers (PLC) are industrial computers whose hardware and software specifically adapted to work in industrial conditions, and that can be easily programmed to build in existing industrial systems. In Figure 2 is shown typical block diagram of the components which is made up of the PLC.

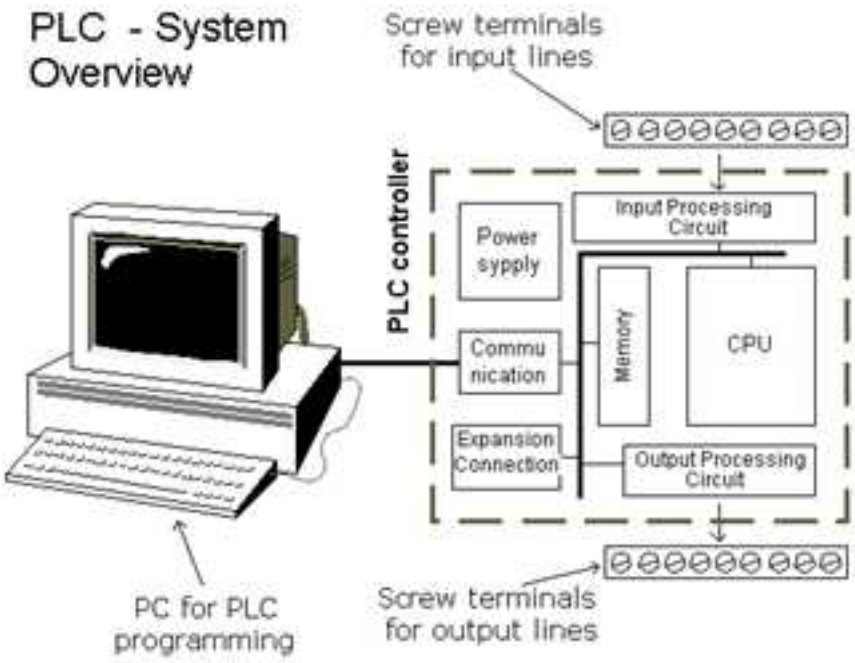

Fig. 2. Cross-sectional Si solar cells

The system that is controllable PLC controller, comprising:

- Input devices such as switches, buttons, sensors, etc.
- The input module, which is part of the PLC. Through this modules receive signals from input devices.
- Logical Unit (or CPU), which represents the "brain"PLC collector consists of a central processing unit and memory. Within this module are saved and the program and data from where it controls the operation of the whole system.
- Output module, which is part of the PLC controller. Through this module selected by the output signals of the individual devices.

- Output devices, such as relay, lamps, motor starters, valves etc.

Particular attention should be paid to the input and output because in these blocks are located and protection necessary for isolating CPU block from the adverse effects the industrial environment via the input lines can be transferred to the CPU. PLC controller is, in essence, a microprocessor device that uses a programmable memory to remember commands that require execution of specific functions such as logic functions, sequencing, counting, timing, calculation, in order to manage different types of machines and processes via the digital and analog input-output modules. PLC has the ability to control its programmable relay (digital) outputs. In fact, the PLC was originally created as a device that replaced the former control relay networks. However, PLC is today much more than just a programmable relay that opens and closes the output contacts. Today PLC represent an efficient and powerful system that can monitor and manage the work of up to several thousand inputs and outputs. PLC can now perform not only the classical management based on the principle of inclusion and exclusion device but a complex control operation of the most complex and diverse actuators, thereby realizing complex technological functions such as PID control, positioning and control wasps etc. One of the significant features of today's PLC represent and their communication capabilities, and networking opportunities and realizing different communication tasks. PLC can also be considered as a special type of computer which is intended to manage various types of industrial processes. Its architecture is very similar to the architecture of standard personal computers, but also specific because it is adapted to its purpose. PLC has a processor and memory, by which is similar to standard PCs, but there are also specific input-output interface by which they are different. PLC has no external memory, but uses only memory that is built into the processor module, which also makes specific computing device. Simpler PLC could be defined as a special industrial computer device that receives signals from the process and based on them will manage certain output devices on the process, and who also has the ability to exchange data with other PLC, PC and other devices.

System for monitoring the apparent movement of the sun

Systems for monitoring the apparent movement of the sun is a pointing device flat collector systems, concentrating solar reflector or lens towards the sun. Concentrators, usually at concentrating systems require a high degree of accuracy and to make sure that the solar radiation is directed exactly to the solar cell, which is at or near the focal point of the reflector or lens. Not concentrating systems do not require a high degree of accuracy, and the monitoring system is not required, but can contribute significantly to the amount of energy produced, improving morning and afternoon performance. Better afternoon performances are especially desirable for photovoltaic systems connected to the network, as production at that time coincides with the peak demand for electricity. Fixed systems are oriented to optimize the limited time it will have a relatively small annual production. For low temperature solar panels, monitoring systems are generally not used, due to the relatively high costs of monitoring system compared to adding a larger collector area and the limited angles of the sun to work in the winter, which affects the average annual capacity of the system. In compared with photovoltaic cells, monitoring systems are relatively cheaper. This makes them particularly effective for PV systems with high efficiency. Most systems to track the apparent movement of the sun in order to most effectively if seasonally adjusts the position of the majority should review and lubricating year.

Control a photovoltaic panel using the PLC

The task of control a photovoltaic panel using the PLC is solved by having the photovoltaic panel set 4 sensors on each side at certain angles and one that is at the center of the solar panels. They also set up two engines, one of which moves the panel down-up and the other left and right. If the central most of the light sensor, the engine will not ignite the solar panel will not move, but as the sun moves after a certain time that the central sensor will be most of the light, but a different one. In this case, the sensor that is was most of the light will turn on a given engine that will move the photovoltaic panel, so that once again the central pod is most illuminated. Figure 3 shows the scheme of SCADA control of photovoltaic panels.

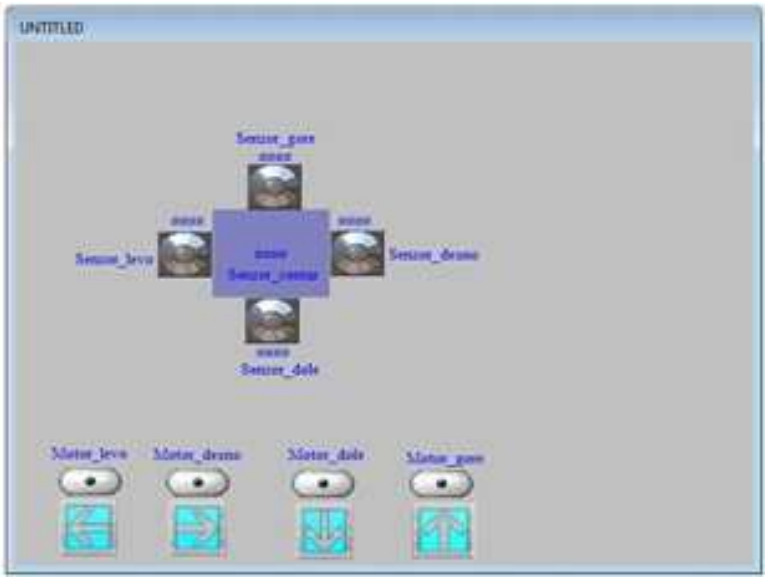

Fig. 3. Scheme of SCADA

Conclusion

Increase of electricity prices in the future will certainly encourage individuals and legal entities to use energy from renewable sources, such as the widely application of photovoltaic panels. For this reason, it is necessary to do everything to create the conditions for market transparency. The legal framework for the production of energy from renewable energy sources, along with incentives, should be directed towards achieving economic, environmental and energy security benefits, to point out the possibility of further investment in plant construction and to stimulate domestic and foreign investors to implement projects relating to the above-mentioned area. The paper presents a model of efficient management photovoltaic panels with the SCADA system and thus enables the efficient production of electricity from solar energy.

References

1. Vera V. Petrovic and Slobodan M. Draskovic, *Control of dynamic systems*, VISER, Belgrade (2015)
2. Petrovic V. and Grujuc A., *Control of photovoltaic panels by using programmable logic controllers* ,ECOLOGICA 21, No. 75, Belgrade (2014).
3. Vera V. Petrovic, *Real time control, manual for programmable logic controllers*, VISER Belgrade (2009)

The Problem of Word Sense Disambiguation in Machine Translation System

D.Rakhimova* and M.Abakan

Laboratory of Intelligent Information Systems,
Institute of Mathematics and Mechanics,
al-Farabi Kazakh National University, Almaty, Kazakhstan
di.diva@mail.ru,mayerabak@gmail.com}

Abstract. This article presents a method for solving the lexical selection problem of nouns in an automatic text processing for different groups of natural languages. The proposed approach is based on micro-context that represented by BoW using feature vectors and NBC for calculating the weight of senses. Wight of senses uses to choose most probable sense of the disambiguated word in sentence. This method has been successfully applied in the machine translation from Russian into Kazakh. The practical results presented.

Keywords: Word Sense Disambiguation, Bag-of-Word, NBC, micro-context, weight, senses.

1 Introduction

The problems of machine translation of texts emerged more than two decades ago. Its essence is to build the machine, on which input enters text in one natural language, and the output generated by a text in another language. Now in the role of the mentioned machine uses electronic computing machines, which are developed algorithms, translators and dictionaries. When translating a sentence from source to target language, a machine translation system has to perform several sub-tasks in order to generate a correct translation. Such a translation should comply to two main properties: faithfulness and fluentness. An important factor in faithfulness is called lexical selection: it is the process of selecting the appropriate translation of source words or phrases amongst the different alternative translation candidates for these words or phrases. It is this process, which described in this deliverable. Fluent translations should be as indistinguishable from sentences originally written in the target language as feasible. An important factor in fluency is the reordering process: the process of appropriately reordering target words and phrases. It would be nice if every word of the language would have one and only one value. But this is not so. When one word has several meanings, they say that there is lexical ambiguity. The fact that ambiguity can accumulate make the analysis very difficult. For example, if the sentence consists of two words and each can have two meanings in the worst case it can be translated in four different ways. Naturally, you may need to consider all possible cases of transfer to select only one of them.

2 Overview of scientific researches and approaches

An important task for machine translation is to find the correct translation of the word in the dictionary, that is, from an existing set of alternative words. Very difficult to find a suitable translation when the number of alternatives increases. This selection plays important role in

* Please note that the LNCS Editorial assumes that all authors have used the western naming convention, with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

quality of translation. Therefore, problem of Word Sense Disambiguation is still a problem in modern machine translation systems. The problem of word sense disambiguation (WSD) has been described as artificial intelligence complete (AI), that is, a problem which can be solved only by first resolving all the difficult problems in AI, such as the representation of common sense and encyclopedic knowledge. Lexical selection problem one of the most important problem of NLP and MT systems. Output of the lexical selection may affect to the sentence translation. It may change the whole meaning of the sentence or may give an incorrect translation. It is possible to distinguish two dominant approaches from the set of all existing algorithms for solving problems of lexical ambiguity. **The first approach** of lexical ambiguity resolution is based on external sources of knowledge (knowledge-based methods). This approach can be easily adapted to the documents obtained from any source and not tied to a specific language.

The second approach is based on machine learning. Algorithms based on this approach show good results in comparison with the algorithms presented in the recent literature, however, they require the training on documents similar to the processed further. This is due to the problem of sparseness of language.

2.1 The Word Sense Disambiguation in Machine Translation System

A special challenge for the translator are ambiguous words. WSD is always a difficult and important task in natural language processing. Its task is to determine the most appropriate sense for an ambiguous word given a context. Approaches for this work include supervised learning, unsupervised learning, and combinations of them. **Micro-context** contains features that that describes disambiguated word. To represent features used BoW, where words represented as some vectors called **features vectors**. Feature vector consisting of numeric values that encodes linguistic information about the language and uses as input for many MT algorithms. This vocabulary (BoW) is pre-selected as some useful subset of words called micro-context. Most of the works on removal of ambiguity use the local or **"micro context"**, as the main information source for determining the correct values. David Yarowsky [1] observed that the length of micro-context may vary depending on the type of polysemy. He suggested that to solve local ambiguity enough 3-4 words of the context, while for semantic ambiguity, the more the window, consisting of 20-50 words. Thus, researchers have still not come to a consensus regarding the optimal length of micro-context. He determined that for words belonging to different parts of speech, the greatest contribution for the solution of lexical ambiguity have different sources. So to solve polysemy of verbs the most important are an objects than subjects, for nouns are the most important related adjectives and nouns, and adjectives for almost all information about their values can be derived from the nouns that they modify. At the moment, to eliminate ambiguity, mainly uses information about the parts of speech of words, in combination with other methods not related to parsing. Yarowsky [2] showed that while a large context can be used to determine the values of the nouns, for solution of lexical polysemy of verbs and adjectives the importance of the context decreases rapidly with increasing distance to the target word. This indicates that for a successful determination of the values of all words in text, you want to use as micro - and thematic context, and to eliminate the ambiguity of different types of words require different methods.

2.2 Description of Bayesian classifier method using the improved BoW method to solve the lexical selection problem

Our proposed **micro-context** consist of most meaningful parts of speech as **nouns, adjectives and verbs**. Let our sentence have one polysemy word W. The word W may have many possible

senses $S = (s_1, s_2, s_3, ..., s_k)$, and the features that describe our context $F = (f_1, f_2, f_3, ..., f_n)$. For example, the polysemy word W="kist (brush in russian)" may be translated into s_1 ="qyl-qalam (brush for drawing in kazakh)" or s_2="bilek (wrist in kazakh)", where $s1, s2 \in S$. By the notion of Yarowsky, we do not need to analyze whole sentence, we just need to find the micro-context of target word. So for sense s_1 its micro-context will be represented as features (words) $F = (f_1, f_2, f_3, ..., f_i)$, and for sense s_2, its micro-context will be represented as $F' = (f'_1, f'_2, f'_3, ..., f'_j)$. The scientific novelty of approach consists in fact that we classify the micro-context by groups of nouns, verbs and adjectives. Improved BoW representation of micro-context came to the form like:

F for s_1:

o $f^n_1, f^n_2, f^n_3, ..., f^n_i$ for noun micro-context;

o $f^v_1, f^v_2, f^v_3, ..., f^v_i$ for verb micro-context;

o $f^a_1, f^a_2, f^a_3, ..., f^a_i$ for adjective micro-context.

F' for s_2:

o $f^{`n}_1, f^{`n}_2, f^{`n}_3, ..., f^{`n}_j$ for noun micro-context;

o $f^{`v}_1, f^{`v}_2, f^{`v}_3, ..., f^{`v}_j$ for verb micro-context;

o $f^{`a}_1, f^{`a}_2, f^{`a}_3, ..., f^{`a}_i$ for adjective micro-context.

MT must give one the best sense $\hat{\mathbf{s}}$ as a result. To choose best sense $\hat{\mathbf{s}}$ from the set of possible senses $\mathbf{S}$, that in our example contain two senses s_1 and s_2, we analyze whole words in micro-context. How the program must choose one of them? It would be the right decision to let them some weight or some coefficient. This numbers we get from the statistic corpora. Here we use the Naïve Bayes Classifier [3].

$$P(s_i) = \frac{(count(s_i, w_j)}{count(w_j)} \tag{1}$$

$$P(f_j|s) = \frac{(count(f_j)}{count(s)} \tag{2}$$

Many scientists use the BoW method with Naïve Bayes Classifier and get some results. Our approach differs from others in that side, that we not only the context that represented by BoW, we decide to use word‘s grammatical, lexical, morphological features, represent them by feature vectors, and collect them into the micro-context. To choose one sense from the possible senses we use the weight, calculated by Naïve Bayes Classifier and data from the corpora that gives statistics of word frequency. We call this method as Method of Naïve Bayes Classifier classifier using the improved BoW method to solve the lexical selection problem. We say improved BoW because BoW divided into 3 categories of micro-context, noun, verb and adjective.

So, lets make a simple algorithm of how it should work:

1. Given the sentence. W is a polysemy word;
2. Determine the possible senses for this W;
3. Micro-context features of each sense represented by BoW method;
4. Feature vector numeric values of each feature, also represented by BoW method;
5. NBC calculate by using corpora statisitics to get a numeric value for feature vectors;
6. To calculate probability for s_1 and s_2 :

$$P(f_i|s_1) = \frac{(c(f_1)}{c(s_1)} \cdot \frac{(c(f_2)}{c(s_1)} \cdot \frac{(c(f_3)}{c(s_1)} \cdots \frac{(c(f_i)}{c(s_1)} \tag{3}$$

$$P(f_i|s_2) = \frac{(c(f_1)}{c(s_2)} \cdot \frac{(c(f_2)}{c(s_2)} \cdot \frac{(c(f_3)}{c(s_2)} \cdots \frac{(c(f_i)}{c(s_2)} \tag{4}$$

7. Then calculate $P(s(_1)$ and $P(s_2)$:

$$P(s_1) = \frac{s_1}{w} \tag{5}$$

$$P(s_2) = \frac{s_2}{w} \tag{6}$$

8. Using general formula choose sense with greater value of probability

$$\hat{s} = arg \max_{s \in S} P(s) \prod_{j=1}^{n} P(f_j|s) \tag{7}$$

$$P(s_1) \cdot P(f_i|s_1) = P(s_2) \cdot P(f_j|s_2) \tag{8}$$

It should be noted that we research only the nouns. To solve the lexical selection problem program should pass some steps:

Step 1: Check all words for disambiguation before translating in a given the sentence. Program reads word by words from left to right the sentence and check them to disambiguation. Text is a set of sentences. Sentence is a set of words. At first, we need to separate sentence from the text. Then, we analyze every word. Words are represented in table form. Noun has linguistic features like gender, declension and polysemy. In addition, words in Source and Target language“ Russian word and its translation into Kazakh are illustrated. Remember that these feature vectors are represented by BoW method. **Step 2**: Determination of senses and micro-context for disambiguated word The two senses are allocated in two tables represent micro-contexts of each sense, it very suitable separate them to the program. Program goes to the both tables. The structures of these two tables are the same. Only the micro-contexts are different, similarly their features vectors of BoW are different too. They also distinguished from each other by the unique ID numbers. **Step 3**: Calculating by NBC for WSD to choose the best sense The third step is about the calculating process. To calculate by NBC we need some statistics from the corpus. The corpus allows to the list of words with frequency. From this table we get a frequency number, and put it on our general formula. Keep in mind that these values may change related to the needed polysemy word. After all stages, compare all the senses and choose the most probable one.

3 Practical results and estimation

This Chapter shows the practical results of machine translation from Russian to Kazakh, that developed on the basis of the proposed methods and algorithms of semantic analysis and synthesis of sentences.

According to the results of machine translation from Russian into Kazakh language, developed program solves the following tasks:

- morphological synthesis word formation of the Kazakh language;
- right selection of endings of Kazakh language;
- correct alignment and forming of the end;
- solution of lexical disambiguation.

Consider the next sentence: “On zharil syroi luk.” (He grilled raw onions).

$$P(f_i|s_1) = \frac{(c(f_1)}{c(s_1)} \cdot \frac{(c(f_2)}{c(s_1)} = \frac{zharit}{piyaz} \cdot \frac{syroi}{piyaz} = \frac{410}{826} \cdot \frac{1426}{826} = 0,84 \tag{9}$$

$$P(f_i|s_2) = \frac{(c(f_1)}{c(s_2)} \quad = \frac{-}{sadaq} = \frac{0}{818} = 0 \tag{10}$$

(example does not contain words from second sense)

$$P(s_1) \cdot P(f_i|s_1) = 0,00144 \cdot 0,84 = 0,0012 \tag{11}$$

$$P(s_2) \cdot P(f_i|s_2) = 0 \tag{12}$$

$0,0012 > 0$, which means that correct translation is the first sense. Note that all frequency of words taken from the corpora. As a result, program choose the first sense “Piyaz”.

For the evaluation of machine translation quality, a technique BLEU is used: a copy of the sample translations corresponding language pair prepared. The text of the source language of the translation sample is fed to the machine translator, It turns the text machine translation in the target language. Received text of machine translation compared with the translation in the sample. The comparison performed by the number of matching words and binary combinations of words.

According to the results of various translators there are major errors:

1. Lexical errors in the formation of the Kazakh language. No proper selection or the lack of finishing in the word.

2. Incorrect alignment and the formation of endings in phrases (especially in this case, the personal and the possessive inflection). Wrong systematization sequence endings.

3. Irregularities in the formation of the syntactic structure of the sentence.

4.The wrong translation of the words. This disadvantage can be explained by the absence of the word or base word in the dictionary translator.

The evaluation of the quality of the translators showed different results. Presumably, if adding to base of vocabulary and rules of analysis and synthesis of text, you can achieve better results in the evaluation of machine translation quality.

In conclusion, one should say that the objective of this analysis is to identify errors or deficiencies of one or another translator. The author wanted to show the effectiveness of the proposed method in the implementation of machine translation, which has relatively good practical results when translating from Russian into Kazakh language.

4 Conclusion

Developed and studied models and algorithms for morphological analysis of Russian words, which are determined by the grammatical characteristics of the words of the Russian language of the original sentence. Algorithms for the analysis of Russian words based on the developed models analysis of words with regard to the grammatical rules of English. Posted derivational forms and types of changes in the number, time, and the case of Russian words. Systematized the end of words. Various properties of the morphology of the Russian language for their transformation into Kazakh language;

Developed and researched models and algorithms of synthesis of the words of the Kazakh language according to the characteristics of Russian words. The developed model of the synthesis of the Kazakh language according to the characteristics of Russian words based on the developed models m-maps of Russian words in grammatical features. Model synthesis of the words of the Kazakh language according to the characteristics of the words of the Russian language developed for all types of parts of speech in both languages.

Developed and studied models and algorithms for lexical polysemy of the words in the sentences of the Kazakh language on the characteristics of offers in English language. The developed model resolve the lexical ambiguity of words on the basis of the context vectors (CV). Method of CV is an unordered set of the most frequent context words generated by processing some body text for the target language. Then for each sense of the polysemy words of the target language a binary vector CV, which consists of elements closest in the meaning of words. The analysis grammars Russian and Kazakh language were defined the basic types of elements many CV, including noun, verb and adjective group. Each group of element values CV, were assigned to the corresponding coefficients of preference. With using the CV elements and their coefficients of preference analyzed the value of the words in the sentence and its conversion into a target language;

Developed a program evaluation of machine translation quality according to the method BLEU.

References

1. David Yarowsky.: Decision lists for lexical ambiguity resolution: application to accent restoration in Spanish and French. Proceedings of the 32nd annual meeting on Association for Computational Linguistics. , Morristown, NJ, USA: Association for Computational Linguistics (1994).
2. David Yarowsky.: One sense per collocation. HLT 93: Proceedings of the workshop on Human Language Technology. Morristown, NJ, USA: Association for Computational Linguistics (1993).
3. Daniel Jurafsky, James H. Martin.: An Introduction to Natural Language Processing, Computational Linguistic, and Speech Recognition. New Jersesy, Second Edition

Immune Network Technology of Complex Objects Control Based on Computing Clusters Using Virtual Machines

G. Samigulina, Z. Samigulina

The Institute of Information and Computational Technologies, Almaty, Kazakhstan
Kazakh National Technical University named after K. Satpayev, Almaty, Kazakhstan
galinasamigulina@mail.ru,zarinasamigulina@mail.ru

Abstract. In this article were designed immune network technology, algorithms and software for the intelligent complex objects control system in the conditions of uncertainty on the basis of computing clusters using virtual machines. Was proposes a component-oriented approach for creating software, was designed author's component in the form of a dynamic library that contains production data processing algorithm based on artificial immune systems.

Keywords: immune network modeling, complex object, virtual computing cluster, component-oriented approach.

1 Introduction

The development of new information technologies and virtual environments creating tools for the production processes modeling led to the possibility of the development of modern systems of automatic complex industrial projects control based on the integration of classical control theory and artificial intelligence methods [2,5]. Artificial intelligence algorithms, such as artificial immune system (AIS) [2], genetic algorithms, neural networks allow to correct the control strategy in the conditions of parameter uncertainty due to the ability to adapt, like live organisms [16].

A widespread application for engineering problems solving got AIS that can be used both alone and in combination with evolutionary algorithms, such as a hybrid systems based on swarm intelligence. [1] To solve the problems of diagnostics AIS are used for data classification. This approach is presented in work [17] as the example of monitoring the state of the planetary gear. The work [2] describes a method of the detection of faults in engineering tasks based on AIS. Data analysis is performed by using principal components to reduce the size of the parameters characterizing the process. Along with this, AIS adaptive approach is used for the safe and reliable operation of the production systems of asynchronous motors [3].

It is relevant to use the AIS in solving problems of computer vision. In work [4] image processing technology for the turning tool movement measurement is described, errors optimization and tool movement calibration is performed. It is possible to use the hybrid systems based on artificial neural networks and AIS for predicting the effectiveness of surface milling parameters, where some characteristics are considered: cutting speed, feed per cutting and the depth of the cut [5].

Today's enterprises are characterized by a huge number of manufacturing data, so timely monitoring and analysis of the messages archives allow to predict operational risks and prevent accidents. In this case, a high reliability requirements and stable operation of all hardware and software are required for automation system.

Software and automation system testing is costly, time-consuming and takes a lot of time [6]. In this regard, the development of technology, allowing delicately to introduce an intelligent control system at the enterprise is relevant.

In this situation, a promising direction of work with virtual machines is very convenient. Computing cluster can be created on the basis of the enterprise computing potential for parallel processing of data streams received from controlled objects. With the help of virtual machines there is implemented a network which consists of several computers, they tested the work of computing cluster and of the distributed immune network control algorithm [12,13], and then connects to the real network. This approach can be applied not only for the dynamic industrial automation systems, but also for a number of other applications [18].

Section 2 describes a formulation of the problem. The description of the algorithmic support of the immune network technology of complex objects control is given in Section 3. Also there was considered a mathematical model of the formal peptide. Immune network technology software and hardware are presented in Section 4 and 5. At the end of the article there is a conclusion.

2 Problem statement

Problem statement is formulated as follows: it is necessary to develop algorithmic support, software and hardware of the immune network technology of complex objects control on the basis of computing clusters using virtual machines.

3 Algorithmic support of the immune network technology of complex objects control

The base of the immune network technology of complex objects control is the approach of artificial immune systems. AIS can create mathematical models of information processing based on the principles implemented by the protein molecules [2]. Such systems should have the properties of self-organization, self-study and the ability to predict.

3.1 Mathematical model of the formal peptide

A basic element of the mathematical model system is the concept of a formal peptide (FP), which is a mathematical abstraction of the principle of the free energy dependence from the spatial form of the polypeptide backbone. [2] Formal peptide is represented by the following ordered quintet:

$$P =< n,\ U, Q,\ V,\ v > \tag{1}$$

1. The total number of protein units is determined by n and should correspond to the condition $n > 0$.
2. A number of torsion angles is represented by the following expression:

$$U = \{\varphi_k, \psi_k\},\ k = 1,\ ...,\ n \tag{2}$$

where $-\pi \leq \varphi_k \leq \pi,\ -\pi \leq \psi_k \leq \pi$;
k - unit number, an amino acid residue;
φ, ψ- rotation angles.

3. A number of quaternions is represented by the following expression:

$$Q = \{Q_0, Q_k\} \tag{3}$$

where quaternions $Q_k = Q_k\{\varphi_k, \psi_k\}$ are determined by the following formula:

$$Q\{\varphi_k, \psi_k\} = q_1 H_0 + q_2 H_1 + q_3 H_2 + q_4 H_3 \tag{4}$$

where H_0, H_1, H_2, H_3 - Pauli matrices;
The resulting quaternion Q_0 is defined as their composition:

$$Q_0 = Q_1\ Q_2\ Q_3, ..., Q_n \tag{5}$$

4. V is a plurality coefficients:

$$V = \{v_{i,j}\},\ i = 1, 2, 3, 4,\ j \geq i \tag{6}$$

5. Function v without index, represents free formal peptide energy and is defined for each element Q_0 by a quadratic form [3]:

$$v = -\sum_{j \geq i} \nu_{ij} q_i q_j \tag{7}$$

where q_i and q_j- configuration of the corresponding i and j formal peptide.

3.2 Multidimensional data processing algorithm based on AIS

For complex objects control there is used the developed immune network technology for multi-dimensional data processing in real time. Here is an algorithm that implements this technology [7].

Algorithm 1.

1. Formation of a database (DB) and knowledge base (KB) which characterize the complex object control (as DB were taken macros of Excel).
2. The classification of the decisions based on the opinions of experts.
3. Normalization of input data.
4. Reduction of uninformative features.
5. The construction of the optimal structure of the artificial immune network. Matrix formation as standards for each class based on on the opinions of experts which will be considered as antibodies: A_1, A_2, A_3,... n, where n - amount of classes.
6. Construction of the optimal structure of the artificial immune network.
7. Learning of AIS.
8. Matrix formation of images which are viewed as antigens: B_1, B_2, B_3, B_4,....,m, where m - amount of images.
9. Determination of the binding energies between formal peptides (antigens and antibodies). The binding energy of two FP in a bilinear form has the following form [13]:

$$w(P, Q) = -\sum_{i,j=1}^{\ell} w_{ij} p_i, q_j \tag{8}$$

Here, the configuration of the first and second FP is presented as:

$$P = \{p_1, p_2, p_3, ..., p_l\}, Q = \{q_1, q_2, q_3, ..., q_l\} \tag{9}$$

The communication matrix $W = \{w_{ij}\}$ in vector-matrix form is represented by the following formula:

$$w(P, Q) = -[P]^T W [Q] \tag{10}$$

where $[P], [Q]-$configuration vectors.

10. Solution to the problem of image recognition based on the determination of the minimum binding energy between the peptides:

$$W_1 = -x_1^T B y_1, W_2 = -x_2^T B y_2, W_3 = -x_3^T B y_3, \\ W_4 = -x_4^T B y_4, \ldots, W_n = -x_n^T B y_n, \tag{11}$$

where T– transposition symbol, n - amount of classes. Native (functional) stacking of the protein chain corresponds to the binding energy minimum, so the minimum value of the binding energy is determined by the class n, to which belongs this image:

$$n : W_n = \min\{W_1, W_2, W_3, W_4, ..., W_n\} \tag{12}$$

11. Evaluation of the energy error in solving the problem of image recognition based on the properties of homologous proteins. Potential averaging is carried out on the homologs:

$$< E_i^* >_G = \frac{\sum_i^G E_i^*}{G} \tag{13}$$

where the symbol $<\ldots>$ means the averaging on the homologs, G - the number of homologous peptides.
Then, calculating energy of the native structure of the average homologs is determined:

$$< E_i^* >_G \approx E_N \tag{14}$$

Z - factors calculation is made. The value of Z - factor is determined by the average number of standard deviations between energy of native structure and energy of a random chain stacking:

$$Z = \frac{E_N - < E >}{< (E - < E >)^2 >^{\frac{1}{2}}} \tag{15}$$

where $E_N \approx < E_i^* >_G$, $< E >$ - the average number of standard deviations on homologs, E - energy of a random chain stacking.
Step 4. Calculation of prediction risk coefficient:

$$K_R(G_i) = |1 - Z_i|, i = \overline{1, n} \tag{16}$$

where n - amount of homologous peptides.
12. Data analysis and forecasting on the basis of AIS.
13. Decision-making based on the obtained data.

4 Software of the immune network technology of complex objects control

Creation of software for complex objects control is impossible without the use of the corresponding system resources. With the development of microprocessor technology and electronics the intelligent technologies can be used also for data searching and processing, for setting the parameters of the mathematical model or for the development of control signals on the basis of image recognition [8]. It is advisable to use modern immune network distributed systems of automatic control on the basis of computing clusters using the technology of virtual machines (VM).

Fig. 1 shows the architecture the implementation of immune network distributed system of complex objects control in a real enterprise network, the computer capacity of which can be used as a computing cluster for parallel data streams processing received from control objects. By using of virtual machines there can be created a network from several computers, they tested the work of the computing cluster, and then this technology is connected to a real network of the enterprise. According to this, the cross-brand compatibility of the equipment is considered to organize the data collection from a real complex of technical means of automation. Therefore immune network technology of complex objects control in the enterprise is implemented in several stages:

1 Stage. The development of component-oriented software in the integrated environment of Visual Studio

2 Stage. Creating a virtual environment for the developed algorithms testing.

3 Stage. Development of computing cluster on the basis of a virtual machine.

4 Stage. Implementation of the tested technology in the real network of the enterprise.

Below the basic steps are decribed.

1. Component-oriented software

Let consider the principles of component-oriented programming. This area involves the development of dynamically extensible carcasses, on the basis of which evolving systems are built. The evolution of the systems is organized through the development of components and their connection to the extension points through the object-oriented interfaces [3]. Components themselves are independent modules designed for multiple use. The advantage of this approach is the dynamic connection of each component and security of the programming, the failed component cannot break the work of others because it is not related with them functionally [3].

In components there are implemented various intelligent algorithms for production data processing in the form of DLL libraries (Dynamic Link Library), which are as a container that contains a code and data for their use in more than one program at a time. With the DLL libraries the program may be divided into modules, the individual components. Immune network technology of data processing for complex objects control based on the algorithm 1 is presented in the form of the author's dynamic library: AisLib.dll (Artificial Immune System Library). This dynamic link library contains the algorithm of artificial immune systems, based on a mathematical implementation of mechanisms of molecular recognition.

The artificial immune systems approach can be used for image recognition, multidimensional data processing, settings the optimal parameters of the regulator [18]. The algorithm is implemented in the application package MATLAB. Then, the initial program code is configured as a dynamic library installation. On the developed software there was received a copyright certificate by the Committee on Intellectual Property Rights MES RK [9]. Each component is connected step by step to an independent application "SmartAnalysis"for data processing complex objects control. Application of component-based approach allows to implement intelligent application capable to integrate a variety of algorithms of artificial intelligence. The modular structure gives an evolutionary component of the application with the ability to custom extensions and phase connections of the necessary components for a specific task. Feature of the application is the ability to integrate with Scada systems that can substantially improve the quality of complex objects control.

2. Creation of a virtual environment for the immune network algorithm testing

For the virtual environment organization there is used the concept of a virtual machine, which is software or hardware system that can simulate the work of a real computer (has its own BIOS, RAM, hard drive and other peripherals). With the help of a virtual machine, there can

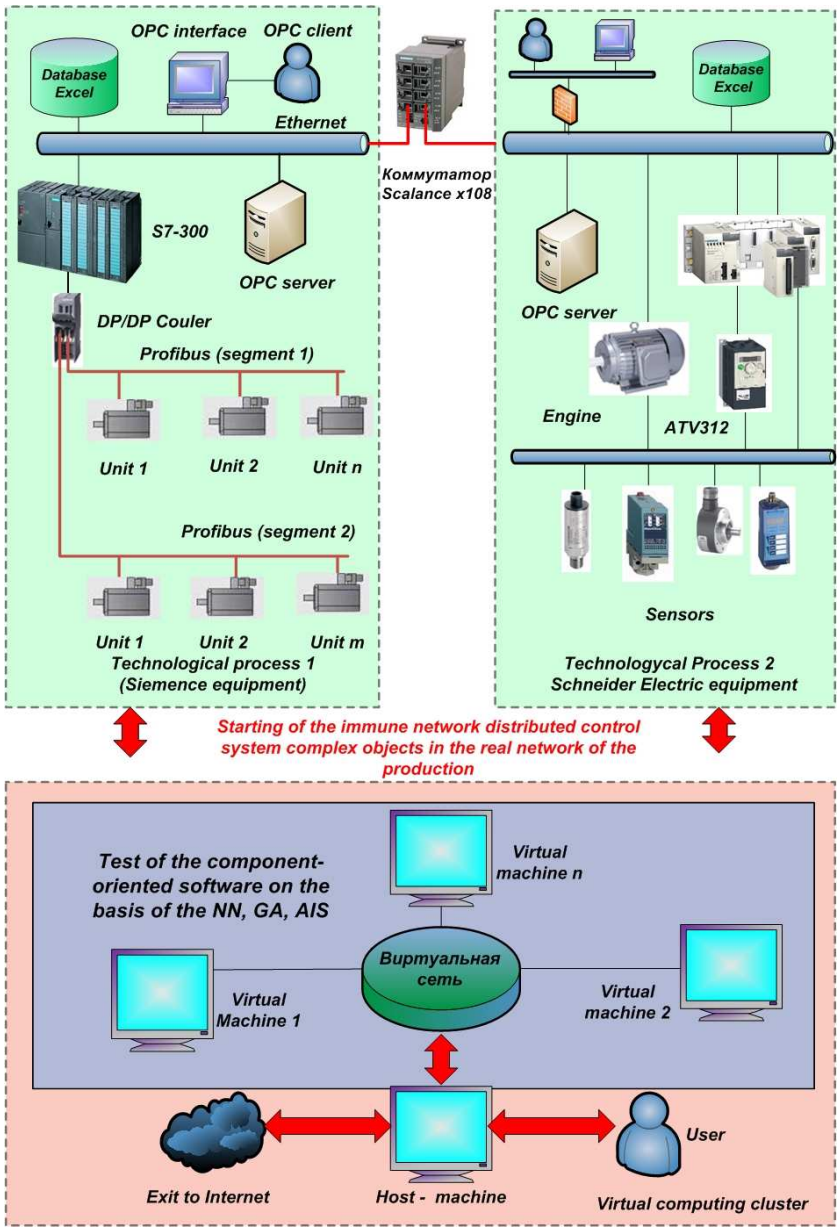

Fig. 1. The architecture the implementation of immune network distributed system of complex objects control in a real enterprise network with the use of virtual machines.

be installed a guest operating system for modeling and various applications. Currently, there are several manufacturers that provide for this work their virtual machines. The most convenient from them is the virtual machine ORACLE VirtualBox with support of Oracle. It allows to install any modern operating systems (Microsoft Windows, Linux, Solaris, BSD, IBM OS / 2, Mac OSx) [12,13].

A real personal computer with a real OS would be called as a host-system and the virtual machine with a virtual OS - the guest OS or the guest system. Creating a virtual machine is made in the application package ORACLE VirtualBox. Then, basic settings of VM configuration are implemented. You must select the operating system platform and version of the installing OS. In our case, the virtual machine will run on the Microsoft Windows platform and OS version

Windows 7. During the use of virtual machines there is allocated amount of memory (RAM). Next, there should be selected the type of a hard disk and a storage format. In our case, we chose a dynamic hard disk. Next, there can be created multiple virtual machines.

The next step is the adjusting of the network between them. On the desktop computer we can create up to five virtual machines connected to a local network, depending on the power source and the characteristics of host-machines. During the setting of the network there can be supported up to four network adapters. Each network adapter has its own functionality: NAT (Network Address Translation), Bridge Adapter (Network Bridge), Internal Network (internal network), Host-only adapter (virtual host adapter) [12,13].

In order to solve this problem it is preferred the configuration of the virtual host adapter that supports the access of VM to a global network, VM is connected over a local network between multiple VM and host-machine. As on the real computer, each virtual machine gets its IP address and subnet mask. Subnet mask, gateway, and DNS servers must match the address set for the network bridge host-system.

After the organization of the virtual network there are created optical bootable disks of MATLAB software in order to create a computing cluster. For this purpose there are used the programs in order to mount ".iso"images (WinCDEmu, ImgBurn, Diamond Tools Lite). After installing of MATLAB on all running virtual machines, you must create a computing cluster [19].

3. Development of a computing cluster on the basis of the virtual machine.

Under the computing cluster is understood a number of computers connected to the local network for specific applications. For this the technologies of Distributed Computing Toolbox (DCT) and MATLAB Distributed Computing Engine (MDCE) are involved on each virtual machine. Creating of a computing cluster using MATLAB is made by the algorithm 3.

Algorithm 3.

Step 1. Installation of MDCE service on each machine.

Step 2. Running a special task manager (jobmanager), whose task is to control the parallel computing.

Step 3. Starting a remote session MATLAB (workers) for the tasks (task).

Step 4. Viewing a created cluster (nodestatus).

Step 5. Search in MATLAB for the created task manager.

Step 6. Getting a link for the object jobmanager in order to send tasks to the account [19].

Step 7. Calculation of the task.

Step 8. Testing the computing cluster.

As a result of the organization of virtual computing cluster the system will solve parallel tasks during the implementation of immune network control algorithms. Using the resources of the enterprise in order to create their computing cluster is a convenient solution in terms of cost savings and optimization of production costs, and the testing and implementation of modern technologies of immune network control using virtual machines increases the level of secure implementation of the immune network distributed control system.

5 Immune network technology of complex objects control hardware

A feature of modern automation systems is the availability of equipment from different manufacturers in the control loop. Organization of cross-brand compatibility, which allows to create a single information database of parameters characterizing the process.

A key part of industrial automation systems is a programmable logic controller for working in local and distributed control systems in a real time. Fig. 2 shows the architecture of a distributed hierarchical combined control system based on equipment of Schneider Electric and Siemens.

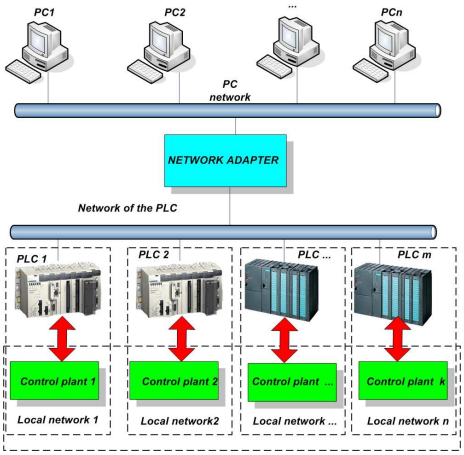

Fig. 2. The architecture of a distributed hierarchical combined control system.

As a process manager there is used on the one hand the PLC Modicon M340 (Schneider Electric), on the other hand Simantic S-300 (Siemens). It is possible to connects data of the controllers directly through the device Scalance x108 (Fig. 1), or with the help of technology OPC [9,21].

From the control object the information is collected in the form of signals from the sensors and fed to the controller, where it is processed. The controller on the basis of the obtained information can adjust the state of the control object acting on actuators. As a data warehouse the Excel tables are used. Organization of cross-platform connection of the equipment from different manufacturers and data processing with the help of artificial immune systems can develop a system of automatic control of the technological processes of the new generation.

6 Conclusion

As a result of the research there were obtained following main results:

- Developed immune network technology of production data processing based on the AIS approach. The advantage of this approach is the estimation of energy error based on the homologous proteins that can recognize FP similar in their structure and at the boundaries of nonlinearly separated classes.

- Presented immune network technology is implemented using component-oriented software in the form of author's components, realized with the help of DLL library AisLib.dll.

- Developed the principles of software testing based on AIS by creating a virtual environment that allows to implement immune network technology delicately in a real enterprise network.

These researches are conducted in the frames of the grant of the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan on the theme: "Development of information technologies, algorithms, software and hardware for intelligent complex objects control systems under conditions of unknown parameters"(2015-2017).

References

1. Elleithy, K., Sobh, T., Iskander, M., Kapila, V., Karim, M.A., Mahmood, A.: Technological Developments in Networking. Education and Automation, P. 643. Springer (2010)
2. Kim Sungshin, Jung Jin-Woo, Kubota Naoyuki: Soft Computing in Intelligent Control. P. 97. Springer (2014)

3. Tarakanov, A.O.: Formal Peptide as a Basic Agent of Immune Networks: From Natural Prototipe to Mathematical Theory and Applications. I Int. Workshop of Central and Easters Europe on Multi-Agent Systems, pp. 123-127. St. Petersburg (1999)
4. Suresh L. Padma, Dash Subhransu Sekhar, Panigrahi Bijaya Ketan: Artificial Intelligence and Evolutionary Algorithms in Engineering System. P. 873. Springer (2015)
5. David F. W. Yap, S. P. Koh, S. K. Tiong, S. K. Prajindra.: A hybrid artificial immune systems for multimodal function optimization and its application in engineering problem. Artificial Intelligence Review, vol. 38, pp. 291-301. Springer (2012)
6. Edyta Brzychczy, Piotr Lipinski, Radoslaw Zimroz, Patryk Filipiak.: Artificial Immune Systems for Data Classification in Planetary Gearboxes Condition Monitoring. Advances in Condition Monitoring of Machinery in Non-Stationary Operations, pp. 235-247. Springer (2014)
7. Chunliu Xiong, Yuhong Zhao, Wei Liu.: Fault Detection Method Based on Artificial Immune System for Complicated Process. Computational Intelligence, vol. 26, pp. 625-630. Springer (2006)
8. Prasant Kumar Mahapatra, Spardha Sethi, Amod Kumar.: Comparison of Artificial Immune System and Particle Swarm Optimization Techniques for Error Optimization of Machine Vision Based Tool Movements. Journal of The Institution of Engineers. Springer (2015)
9. Ramezan Ali Mahdavinejad, Navid Khani, Mir Masoud Seyyed Fakhrabadi.: Optimization of milling parameters using artificial neural network and artificial immune system. Journal of Mechanical Science and Technology, vol. 26, pp. 4097-4104. Springer (2012)
10. Ilhan Aydin, Mehmet Karakose, Erhan Akin.: An adaptive artificial immune system for fault classification. Journal of Intelligent Manufacturing, vol. 23, pp. 1489 - 1499. Springer (2012)
11. Le Thi My Hanh, Nguyen Thanh Binh, Khuat Thanh Tung.: A Novel Test Data Generation Approach Based Upon Mutation Testing by Using Artificial. Knowledge and Systems Engineering, vol. 326, pp. 169-181. Springer (2015)
12. Jones, T.: Application virtualization, past and future. An introduction to application virtualization, IBM Corporation [Electron. resource], URL: http://ibm.com (2011)
13. Andrew Citron, Gaurav Bhagat, Marshall Massengill.: Compare performance of virtual agents with real ones. IBM Corporation [Electron. resource], URL: http://ibm.com (2012)
14. Samigulina, G.A., Samigulina, Z.I.: Intellectual systems of forecasting and control of complex objects based on artificial immune systems. P. 174. Yelm, WA, USA: Science Book Publishing House (2014)
15. Samigulina, G.A.: Development of the decision support systems on the basis of the intellectual technology of the artificial immune systems. Automatic and remould control, vol. 74, pp. 397-403. Springer (2012)
16. Taraq Hussain Sheakh.: Software reliability analysis - a new approach. International Journal of Scientific Research, vol. 3, pp. 75-77 (2012)
17. Ermakov, I.E.: The current system software for component-oriented programming. Proceedings of the II International scientific and practical conference "The Object systems-2010"(2010)
18. Samigulina, G.A., Samigulina, Z.I.: Intellectualization of data processing based on the approach of artificial immune systems for industrial automation. Copyright certificate of the Committee on Intellectual Property Rights MJ RK on software "Data Analizer (2013)
19. Olenev, N.N., Pechenkin, R.W., Chernetsov, A.M.: Parallel programming in MATLAB and in its applications, pp. 43-48 (2007)
20. Samigulina, G.A., Samigulina, Z.I.: Word Implementation of Intellectual Immune Network Technology Controlling the Complex Objects. IEEE 8th International Conference on Application of Information and Communication Technologies, AICT2014, pp. 344-347. Astana (2014)
21. Samigulina, G.A.: Immune network modeling technology for complex objects intellectual control and forecasting system. Monograph, P. 172. Yelm, WA, USA: Science Book Publishing House (2015)

Investigation of Artificial Immune System using Fuzzy Logic

Olga Shiryayeva, Tatyana Denisova

Institute of Information and Computational Technologies of Ministry of Education and Science Republic of Kazakhstan,
Pushkin st., 125, 050010,
Almaty, Kazakhstan
oshiryayeva@gmail.com,elkaz41@mail.ru

Abstract. This article presents the results of the development of immune model of organism reactions to therapeutic doses of drugs based on fuzzy logic methods. There was developed an adequate mathematical model describing the body's response to drugs in the process of infection which is used for developing optimal dynamics of the number of infected, disinfected cells depending on certain initial conditions and the drug dosage described by fuzzy sets. Also, based on the theory of fuzzy sets there were presented: scenarios of infection extension in the body, depending on micro-organisms extension in accordance with the concepts and basics of microbiology; forms of the body infection; change of pharmacological response, depending on the drug doses based on interaction of the body and drugs in pharmacology; unfavorable drug reaction.

Keywords: Artificial immune system (AIS), the therapeutic drug dosage, modeling of the body reaction, the theory of fuzzy sets.

1 Introduction

Today, the immune system is considered by the researchers as a source of ideas and methods of solving various problems in the analysis of biological systems with population dynamics, information processing and analysis, mathematical modeling and information security [1]. Currently, the number of works on the development and application of artificial immune system (AIS) increased rapidly. Individual articles on AIS began to appear in the 1980's, but separate areas of AIS appeared only in the middle of 90s with the appearance of Forrest, Dasgupta, Hunt, Cooke works [2]. Every year since 2002, there is held an international scientific conference dedicated to this field of research (International Conference on Artificial Immune Systems, ICARIS).

The immune system of organism is a complex adaptive structure that effectively uses a variety of mechanisms of protection against external pathogens [2]. The main task of the immune system is the recognition of cells (or molecules) of the body and their classification as their own or others. Detectable foreign cells serve as a signal for the activation of the protective mechanism of the appropriate type.

The immune system - is a structure in which the mechanisms of learning, memory, and associative search are implemented to solve the problems of recognition and classification [2]. In particular, the immune system can be trained to recognize the important structures (antigenic peptides); for memorizing of already encountered structures and for the use of the laws of combinatorics within gene libraries for the efficient generation of structure detectors (variable regions of antibody molecules) interacting with external antigens and the body's own cells. However, the reaction of the antigen takes place not only at the level of individual unit recognition but also at the system level through mutual recognition of lymphocyte clones in antigen-antibody reactions. Thus, the behavior of the immune system is determined by the totality of the local network interactions. Immune system attracts a great interest because of its important role in

maintaining the integrity of the organism. The properties of the immune system are a remarkable example of local adaptive processes, implementing effective global responses [3].

Currently, under the conditions of the complex dynamics of the body's reaction to drugs the development of modeling procedures and, in the future, creating the optimal response of the body for achieving the complex of variety of purposes related to the measure of the effectiveness of the body protection, which has the plurality of evaluation criteria, is becoming one of the most important tasks of medicine [5]. In connection with this the problems of structure creating of optimal immune system effect of drugs on the human body, where the control can be seen as a function of time, reflecting the possible impact on the disease treatment process are under a big interest. During the setting the targets for creating an optimal structure of the immune system influence of drugs on the human body in case of such complex phenomena as the processes of the body, it is quite difficult to control adequately and to construct a satisfactory measure of the quality achievement of the complex of a variety of goals that will improve the body condition, which is one of the problem for study of which it is relevant to develop specialized techniques with algorithmic programming and software.

The developing of immune model of the body reaction to drugs based on the methods of artificial immune systems should take into account the uncertainty in the parameters description, due to the nature of processes occurring in the body. Nowadays, there are various methods of uncertainty representation, also on the basis of the theory of fuzzy sets. This is conditioned by the performed analytical review of the existing methods of representation and research of fuzzy immune systems, on the basis of which there was made a conclusion about the prospects of this direction [2]-[9].

Today, the basic laws and mechanisms of immune system dynamics are used to evaluate and forecast the dynamics of populations of immune cells in a form of controlled mathematical models which allow to explore the protective mechanisms of the organism to the influence of external antigens. These results allow to obtain a methodology of quantitative evaluation of a therapeutic value of innovative medicines. In particular, the methodology for developing the treatment programs for complex immune diseases such as HIV-infections, improving treatment results through the use of mathematical technologies while reducing the volume of drugs [5]-[7]; treatment of gliomas with the solving the problem of an optimal therapy strategy searching, i.e. a search of the dose and time of taking the medicine when to the final time the total number of glioma cells was minimal [8].

Based on the above statement the problem is composed as follows: on the basis of a mathematical model of artificial immune system of an infectious organism disease, developed by the author in works [9],[15] to get the results of constructing an immune mathematical model structure simulating different algorithms of formation of special organism reactions, depending on the strategies of infection based on the theory of fuzzy sets.

2 Development of artificial immune system based on fuzzy logic

Let the mathematical model of infectious organism disease system and the influence on it of the drugs presented by the system of equations connecting the dynamics of change in the number of cured cells by the action of drugs, and the number of infected cells by the action of the disease and unfavorable effect of drugs and disease to body resources has the form [9]:

$$\begin{cases} \frac{dL(t)}{dt} = (-\gamma - nP(t) - gC(t))\, L(t) \\ \frac{dP(t)}{dt} = (-\beta + \lambda(1/L(t)) - \rho C(t))\, P(t) \\ \frac{dC(t)}{dt} = (-\alpha + \mu P(t))\, C(t) \end{cases} \quad (1)$$

where L – quantitative representation of the body resources; P – the number of infected cells of the organism depending on time; C – the number of detected and disinfected cells using the drug, evaluates the effectiveness of the drug; γ – the rate of natural increase of the resources of the body; n – the amount of resources for one infected cell; g – the amount of body resources to respond to the drug, evaluates the unfavorable effect of the drug; β – the rate of change in the number of infected cells in the absence of body resources; $-\beta+\lambda$ – the cascade increase in the number of infected cells with the help of the resources; ρ – fuzzy value of the detection of the infected cell by the drug; α – the rate of change of the quantity of the detected cells based on time and the volume of the action of the drug in the absence of infected cells; μ – the rate of cascade change in the number of the detected and disinfected cells with the occurrence of viruses.

Interference of the variables $L(t),\ P(t),\ C(t)$ is defined by the parameters $n,\ g,\ \lambda,\ \rho,\ \mu$. The parameters $\alpha,\ \beta,\ \gamma$ define own dynamic of the variables $L(t), P(t), C(t)$. In order to determine the nature of the mathematical model (1) and parameter values based on the theory of fuzzy sets it is necessary to consider the processes of infection extension in the body, the effect of drugs on the infection and the body's resources for solving the problems of forecasting and development of effective influence of drugs on the body with minimizing of the unfavorable effects.

In work [16] there are widely presented the results of the immunological models construction of the various stages of infectious diseases. Various forms of diseases associated with different stages of infection of the organism, the interaction of organism and infections, types and forms of infections. Based on the classification of the forms of the organism infection [16] a mathematical model of infectious disease of the organism (1) describes the step of primary manifestations, subclinical stage and secondary stage of the disease, over the μ coefficient of the equation, the number of detected and disinfected cells.

In accordance with the concepts presented in [16]-[10] let make a Table 1 of the immune mathematical model parameters (1). To do this, there were considered the following factors:

- scenario of infection extension in the body, depending on the microorganisms extension in accordance with the concepts and basics of microbiology [16];

- forms of the body infection [17];

- change of pharmacologic response, depending on the doses of drugs based on interaction of the body and drugs in pharmacology [18];

- unfavorable reactions [10].

In the developed mathematical models there is considered a therapeutic use and a molecular level of drug effects.

In accordance with the process of implementation of the evaluation parameters of the mathematical model of the immune system and programming tools of MATLAB there were obtained diagrams of different formation algorithms of special reactions of the organism to drugs, depending on the strategy of the reactions of the body infection in accordance with different scenarios shown in Table 1.

Table 1. Parameters of the mathematical model (1)

$\frac{dC(t)}{dt} = (-\alpha + \mu P(t)) C(t)$	
$\mu = \mu_P$	coefficient of the cascade change in the number of detected and cured cells on the stage of primary symptoms
$\mu = \mu_A$	coefficient of the cascade change in the number of detected and cured cells with the occurrence of the viruses of subclinical stage
$\mu = \mu_S$	coefficient of the cascade change in the number of detected and cured cells on the stage of secondary symptoms
α	coefficient of change of the quantity of the detected cells based on the duration and action volume of the drug in the absence of infected cells
α=α_{low}	low coefficient of change of the quantity of the detected cells with the local infection is different by the accumulation of microorganisms in the hearth of infection
$\frac{dP(t)}{dt} = (-\beta + \lambda(1/L(t)) - \rho C(t)) P(t)$	
$\rho = \rho_1$	Threshold drug dose
$\rho = \rho_2$	Average drug dose
$\rho = \rho_3$	Highest drug dose
$\rho = \rho_4$	Threshold lethal dose

In accordance with the process of implementation of the evaluation parameters of the mathematical model of the immune system and programming tools of MATLAB there were obtained the diagrams of different formation algorithms of special organism reactions to drugs, depending on the reactions of the body infection strategies according to different scenarios shown in Table 1. The initial conditions for variables, parameters and constants of the mathematical model (1) are shown in Table 2. The simulation scheme in MATLAB is based on a discrete mathematical model.

Table 2. Variables and parameters for the model simulation (1)

L_0	P_0	C_0	α	β	γ	μ	ρ	n	k	λ
0.01	10	1	-0.1	-0.1	-0.1	0.1	1	200	200	0.2

The parameters for the mathematical immune model of the organism response to drugs on the stage of primary symptoms (Fig. 1) are given in Table 1.

The parameters for the mathematical immune model on the subclinical stage (Fig. 2) are given in Table 1 and are determined by the parameter $\mu_A = 0.2$.

The parameters for the mathematical immune model with the local infection (Fig. 3) are given in Table 1 and are determined by the parameter $\alpha_{low} = -0.0001$.

The parameters for the mathematical immune model with the highest drug dose (Fig. 4) are given in Table 1 and are determined by the parameter $\rho_3 = 5$.

The parameters for the mathematical immune model with threshold lethal dose (Fig. 5) are given in Table 1 and are determined by the parameter $\rho_4 = 10$.

On the basis of the constructed mathematical model (1) and obtained simulation results (Fig. 1-5) there can be analyzed the reactions of the organism to virus and drugs.

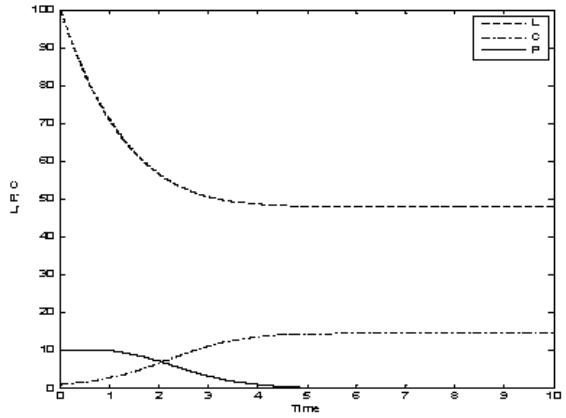

Fig. 1. Results of immune model simulation of organism reaction to drugs (1) on the stage of primary symptoms

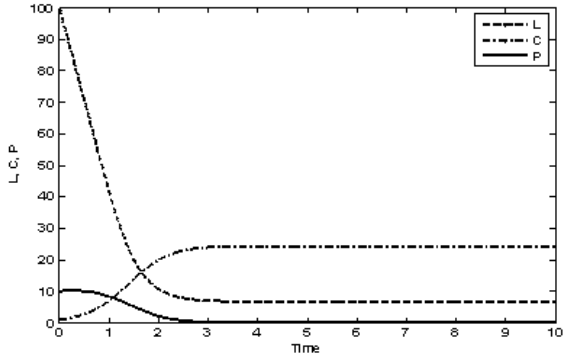

Fig. 2. Results of immune model simulation of organism reaction to drugs (1) on subclinical stage

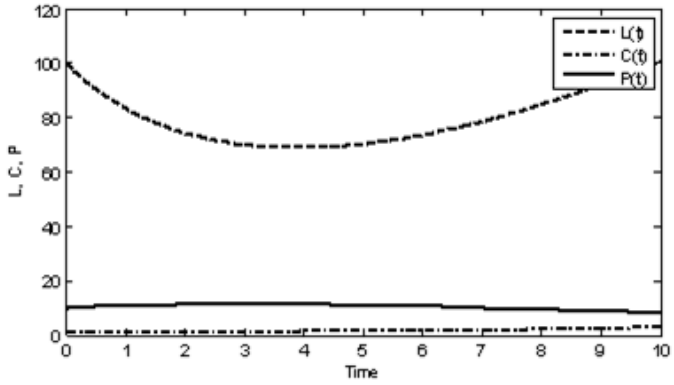

Fig. 3. Results of immune model simulation (1)with the local infection

Let the resources of the body is sufficient for the occurrence of infectious processes. In this case, after the organism infection similar to the body's immune system, a period of latent disease occurs – latent period when the virus penetrated into the organism infects cells without the

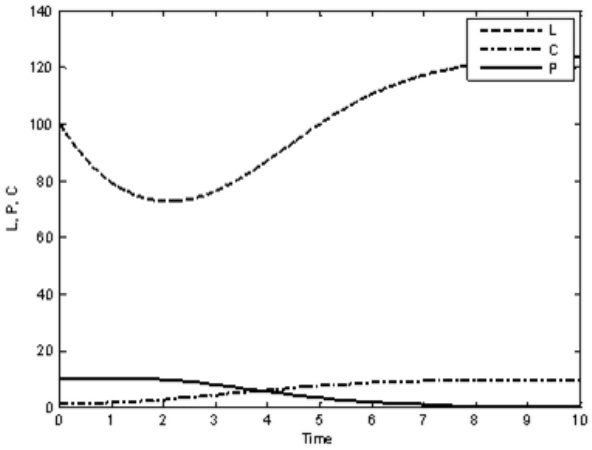

Fig. 4. Results of immune model simulation (1) with the highest drug dose

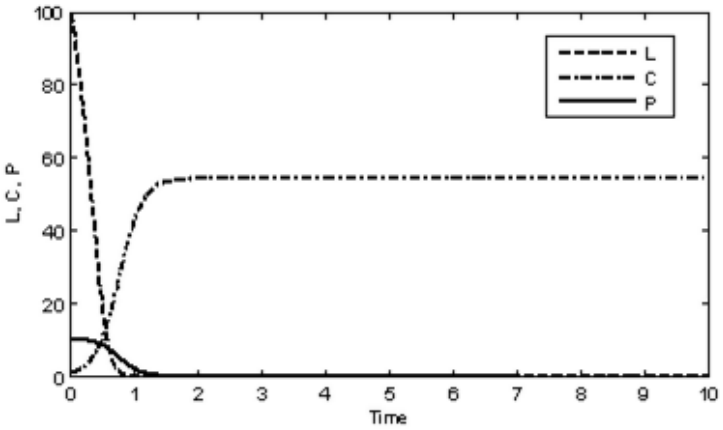

Fig. 5. Results of immune model simulation (1) with threshold lethal dose

occurrence from the immune system a sufficiently severe reaction. If on the step of primary symptoms (Fig. 1) the drug dose enters the organism then starts a complete organism recovery. However, the process of the virus recognition at this stage makes its usage difficult.

After appearing the first detectable response to infection in the organism, the subclinical stage begins – the stage of growth. At this moment the most active immune reaction to the virus starts. During this stage, time for virus recognition is higher and drug dose entering as the threshold dose results to a complete recovery of the organism.

Fig. 1 and Fig. 2 shows the results of the modeling of all classification scenarios (Table 1) of infection developing in the organism, depending on the extension of microorganisms when microorganisms or products of their activities are distributed throughout the organism. Fig. 3 is considered to the organism response to drugs at a local infection, when a low coefficient of change of the detected cells α_{low}= -0.0001. At the same time, a high drug dose coming into the organism leads to a full recovery, but it takes a critical amount of organism resources (Fig. 4). In the case where the drug dose exceeds the threshold lethal dose, the organism resources are not sufficient to cure (Fig. 5).

Developed immune model (1) of the organism reaction to drugs based on the methods of artificial immune systems is nonlinear and stochastic. Fuzzy models of the systems are universal

approximators of a wide class of nonlinear and stochastic functions, allowing due to the presence of a training sample to approximate the function or the measured data with the required accuracy. On the other hand, fuzzy immune models allow to describe the phenomena and processes occurring in the human body, in natural language by means of linguistic variables.

One of the characteristic of a fuzzy set is a belonging function. The reality of the disease is estimated with a certain degree. We denote $\mu_A(x)$ - the belonging degree to the fuzzy sets A, which is a generalization of the concept of the characteristic features of conventional sets. Then, a fuzzy set A is a set of ordered pairs of the type $A = \mu_A(x)/x$ [0,1]. The value $\mu_A(x) = 0$ means the absence of belonging to the set, 1 - full belonging. For the fuzzy sets, as for ordinary ones, there are defined logical operations, the main of which are negation, intersection and combination. Belonging functions can be defined by the various ways. Thus, for example, there are used in practice triangular, trapezoidal, Gaussian and other belonging functions. Fig. 6 and Fig. 7 shows a representation of linguistic variables by the triangular functions, according to the model (1) and to the Tables 1, 2:

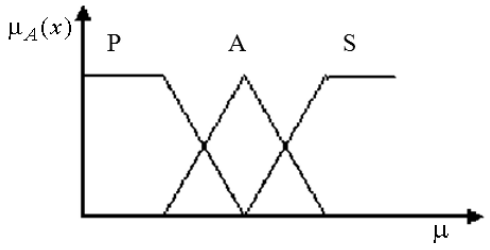

Fig. 6. Belonging function $\mu_A(x)$ of the cascade change in the number of detected and cured cells

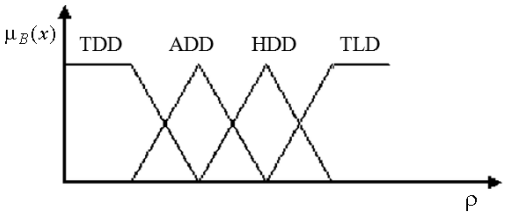

Fig. 7. Belonging function $\mu_B(x)$ of the drug dose

P – the cascade change in the number of detected and cured cells on the stage of primary symptoms;

A – the cascade change in the number of detected and cured cells with the occurrence of the viruses of subclinical stage;

S – the cascade change in the number of detected and cured cells on the stage of secondary symptoms;

TDD – threshold drug dose;

ADD – average drug dose;

HDD – highest drug dose;

TLD – threshold lethal dose.

The parameters for the mathematical immune model of the organism response to drugs given in Table 1 can be described as fuzzy sets $\mu_A(x)$and $\mu_B(x)$.

In the work is used the fuzzy conclusion on the example of Mamdani mechanism. This is the most common way of logical inference in fuzzy systems. It uses minimax formulation of fuzzy sets. Classic fuzzy systems have the disadvantage in which for the rules and belonging functions formulation it is necessary to involve experts of a particular area that is not always possible to provide. Adaptive fuzzy systems solve this problem. In such systems, the selection of the parameters of the rules of the fuzzy output is produced in the process of training on the experimental data. Learning algorithms of the adaptive fuzzy systems are relatively time consuming and usually consist of two stages: the generation of linguistic rules and adjusting the parameters of belonging function and the structure of the rules. The first task relates to the problem of the exhaustive search type, the second – to the optimization in continuous space. A significant part of fuzzy systems learning methods are used the genetic algorithms (GA). However, the classical GA and their varieties are not always effective in solving the problem of multimodal optimization. The development of evolutionary computation to improve the efficiency of fuzzy models learning algorithms is conducted towards the creation of new methods using the possibility of dynamically change of the parameters of optimization algorithms. AIS, to which such properties as pattern recognition, diversity, training, methodinamics and others are inherent, meets to such requirements. In work [3] is shown the application of AIS in fuzzy classification systems, and also for the generation of fuzzy rule base.

The main task - to form a plurality of fuzzy rules that detect anomalies, constructed from data that describe normal and anomalous system behavior. The method of finding the set of fuzzy rules to determine the anomalous behavior of the system is based on the use of multi-dimensional data containing information about parameter values of the network traffic. Problem statement. The purpose of anomalies detection is to identify the current state of the system, which may be normal or anomalous.

3 Conclusion

In the article was developed a mathematical model describing the organism response to drugs in the process of infection used to develop the optimal dynamics of the number of infected and cured cells depending on the relative initial conditions and the appropriate therapeutic dose, based on fuzzy logic methods.

For the parameters of the mathematical immune model of the organism reactions to drugs there were generated fuzzy scenarios of infection extension in the organism, according to the forms of organism infection and to the therapeutic doses of drugs: fuzzy set of cascading changes in the number of detected and cured cells, fuzzy set of therapeutic doses of drugs.

Conclusions: The developed mathematical model of the drugs influence on the human body based on the artificial immune systems approach, gives an adequate tool for the study of the processes associated with the organism response to drugs in the process of infection and will be used to develop optimal dynamics of the number of infected and cured cells depending on relative initial conditions and drug doses.

Work supported by the grant of Ministry of Education and Science of Kazakhstan Republic N0115PK00549: "Computer-aided molecular drug design based on immune network simulation".

References

1. Velichenko, V.V., Pritykin, D.A.: Of artificial intelligence and computer technologies for developing treatment programs for complex immune diseases. Fundamental and Applied Mathematics, vol. 15, pp. 21-42 (2009)

2. Dasgupta, D.: Artficial Immune Systems and Their Applications. P. 438. Springer-Verlag, New York (1998)
3. Webb, G.F., Kirschner, D.E.: Using mathematics to understand HIV immune dynamics. Notices Amer. Math. Soc., vol. 43, pp. 191-202 (1996)
4. Webb, G.F., Kirschner, D.E.: Resistance, remission, and qualitative differences in HIV chemotherapy. Emerging Infectious Diseases, vol. 3 (1997)
5. Murray, J.D.: Mathematical Biology II: Spatial Models and Biomedical Applications, P. 150. Springer, New York (2003)
6. Nabil, E., Badr, A., Farag, I., Khozium, M.: A Hybrid Artificial Immune Genetic Algorithm with Fuzzy Rules for Breast Cancer Diagnosis. INFOS2008, Cairo-Egypt (2008)
7. Polat, K., Sahan, S., Gune, S.: A new method to medical diagnosis: artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia. Expert Systems with Applications, pp. 264-269 (2006)
8. Shamshirband, S., Hessam, S., Javidnia, H., Amiribesheli, M., Vahdat, S., Petkovic', D., Gani, A.: Kiah MLM. Int. J. Med. Sci, pp. 508-514 (2014)
9. Polat, K., Kara, S., Gunes, S.: A novel approach to resource allocation mechanism in artificial immune recognition system: fuzzy resource allocation mechanism and application to diagnosis of atherosclerosis disease. ICARIS, pp. 244-255. Springer-Verlag, Berlin (2006)
10. Velichenko, V.V., Pritykin, D.A.: Office of AIDS treatment. Automation and Control, pp. 166-185 (2004)
11. Velichenko, V.V., Pritykin, D.A.: Numerical methods of optimal control the dynamics of HIV. Math. Russian Academy of Sciences. Theory and control systems, pp. 53-64 (2006)
12. Joshi, H. R.: Optimal control of an HIV immunology model. Optimal Control Appl. Methods, vol. 23, pp. 199-213 (2002)
13. Kovalenko, S.J., Bratus, A.S.: Estimates of the optimality criterion in the problem of modeling therapy gliomas. Mathematical Biology and Bioinformatics, vol. 9, pp. 20-32 (2014)
14. Shiryayeva, O.I.: Development of a mathematical model of the immune effects of different scenarios of medicine. KBTU, Almaty (2013)
15. Shiryayeva O.I.: Terms of constructing a mathematical model of the immune effects of medications on the human. Scientific forum dedicated to the 75th anniversary of professor Biyasheva, pp. 224-228. IPIU, Almaty (2013)
16. Marchuk, G.I.: Mathematical models in immunology. Computational methods and experiments, P. 304. Nauka, Moscow (1991)
17. Zhmakin, A.I, Shejbak, V.M., Ostrovcova, S.A.: General microbiology, P. 364. Grodno: GrGMU, Moscow (2012)
18. Gilman. Clinical pharmacology of Goodman and Gilman. P. 400. Practice, Moscow (2006)
19. Viktorov, A.P., Malcev, V.I.: Mechanisms of medicine side effects: problems of terminology and classification. Rational pharmacotherapy, pp. 25-29 (2007)

The Development of Information, Telecommunications, and Computer Technology to Work with Scientific Data in the Russian Far East

Sergey Smagin*, Alexey Sorokin*

* Computing Center of the Far Eastern Branch,
Russian Academy of Science, Khabarovsk, Russia
alsor@febras.net,smagin@ccfebras.ru

Abstract. Over the past 10 years, the formation of basic telecommunications and computing infrastructure has been completed in the Far Eastern branch of Russian Academy of Sciences, which provides an opportunity to scientists to do their work using the most modern information technology.
The results obtained allowed to proceed to the implementation of projects related to the construction of distributed information systems for large-scale integrated research in the geosciences.
Further development of telecommunication systems in the Far East of Russia and an increase of network channel capacity provides for the future creation of information systems and services of the new level, meaning the formation of the common scientific information environment in the region.

Keywords: data network, communication channel, information system, information technology, instrumental data, data center, data processing.

1 Introduction

A number of scientific institutions located in the Russian Far East conduct basic and applied research on the most important problems of natural, technical, social and humanitarian sciences. The vast territory, the need to meet the challenges of collecting, processing and exchange of scientific data, remote access and control to the means of instrumental observations and other relevant tasks form the objective needs to use a wide range of information, telecommunication and computing technologies.

One of the tools for solving the above problems is a communication network. In addition to accessing resources of global networks, the network is used to integrate systems and services that operate with geographically distributed sources of information, data storage and processing centers, and so forth.

One of the largest non-commercial networks in the region has been implemented In the Far Eastern Branch of the RAS - Regional Computer Network FEB RAS (Network), connecting 26 scientific institutions of the Federal Agency of Scientific Organizations in Russia. A Data Center, which is developing and implementing IT technologies for working with scientific data, has been established in the Computing Center FEB RAS in Khabarovsk.

The article provides a brief description of Network, as well as examples of developed information services.

1.1 Factors Affecting the Development of Corporate Networks in the Russian Far East

The Russian Far East still has a low and non-uniform level of communication services market. The following main factors affect the cost, availability, and specifications of telecommunications services provided by the service providers:

1. The lack of ground communication links with the northern regions (Magadan region, Kamchatka region, Chukotka Autonomous District). These destinations implement satellite channels that have a delay of the data transmission about 600 ms, there are also limitations in the data transmission capacity and high-speed capabilities of the communication channels.
2. Peer connections between the networks of federal service providers, implemented, as a rule, in the exchange sites, located in the European part of Russia. In this case, the data packets have to cross the country twice and numerous transit communication nodes. In addition to the extra time to transfer data, this scheme also generates additional risks associated with the possible components failure of the large and lengthy telecommunications infrastructure used for such a connection.
3. A limited number of service providers who have their own cable backbone data network. This reduces competition and thereby holds the high prices for communication services compared with average level, especially, at the dedicated channels.

These factors affect the quality and quantity of scientific networks implemented in the region. An important step towards minimizing their negative impact at systems development of telecommunications in the Far Eastern branch of Russian Academy of Sciences, was the adoption in 2004 the Program "Information and Telecommunication Resources FEB RAS."

1.2 Regional Computer Network FEB RAS

The main objective of the Program is to implement and maintain IT-infrastructure to provide the necessary types and amounts of hardware, software and telecommunications resources for academic institutions at the modern level of fundamental and applied research.

As part of the program a Network has been implemented, which connects the local networks of Amur, Kamchatka, Primorye (MPPI network FEB RAS), Sakhalin, North-East and Khabarovsk scientific centers. At present, the Network has the following topology and the management structure (Fig.1). All regional scientific centers of the FEB RAS (located in Khabarovsk, Vladivostok, Blagoveshchensk, Petropavlovsk-Kamchatsky, Magadan, Yuzhno-Sakhalinsk) have their regional segments and supporting network nodes which provide LAN routing and operation of the basic network services (DNS, mail, www, proxy, etc.), as well as specialized information systems supported by the local institutions, FEB RAS.

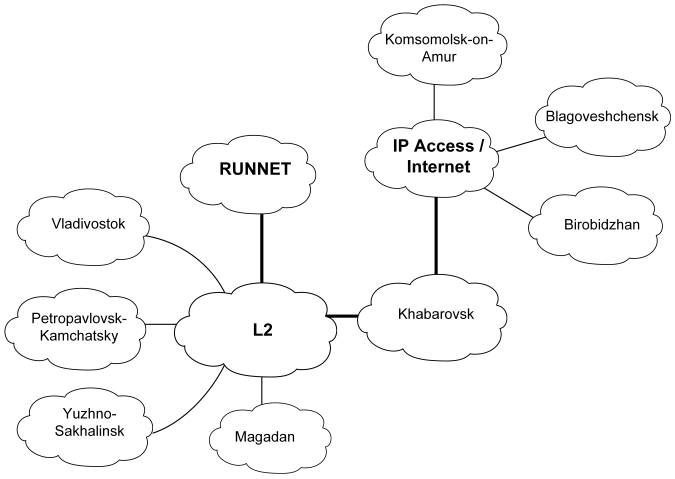

Fig. 1. Network topology.

The Computing Center FEB RAS (CC FEB RAS, Khabarovsk) hosts the Network Control Center (hereinafter - the "NCC"), which performs the following tasks:

- Setup and maintenance of the basic information security system;
- Managing Routing and quality of communication services QoS;
- Support for centralized systems information services (videoconferencing system, IP-telephony, etc.);
- Providing resources for functioning of specialized information systems;
- Advising organization's administrators and users of the Network.

For linking the network segments and accessing the resources of global networks the different types of communication channels and communication technologies are being used:

1. **Dedicated Channels.** Most of the used channels are level L2 dedicated channels, which are used for "point to point"communication, and isolated subnets that provide various services for information systems. This is achieved using IEEE 802.1 Q standard and its expansion Q-in-Q. All the dedicated ground links in the Network are implemented according to this principle.
   To connect Kamchatka and the North-East networks, Far East Branch research locations in remote areas, dedicated satellite links are being used. This requirement is due to guarantee that they receive and transmit data that it is currently impossible to implement at the sites using the standard Internet access service.
2. **Access to the Internet (IP Access).** In the city of Blagoveshchensk, Birobidzhan, Komsomolsk-on-Amur, where no applications requiring quality assurance of communication services (QoS) or specific parameters of information systems, the FEB RAS institutions are connected to the Internet through IP Access service. Using the same service provider in all these cities, means that only its local routing nodes are used, that optimize routes and time delay in the interaction of these segments with each other and with the core Network, that is also connected to this service provider via IP-connection.
   A dynamic routing protocol BGP is used to communicate with external networks and IP-based networks announcement. Multiple connections to a global network of Internet provides for means of load balancing mechanism for channels and selection of the best route to the destination network. This provides direct access to the information resources of commercial service providers within the Far East (including network segments connected to it) and to the Russian part of the Internet as a whole.
   An important element of the telecommunications infrastructure of the Network are direct connections with the other scientific and educational networks. Currently, high-speed links to a data networks implemented for Siberian Branch of Russian Academy of Sciences (40 Mbps L2 to Novosibirsk from Khabarovsk) and the regional segment of the Federal Russian University Computer Network RUNNET (1 Gbps).

2 Network Management System

The heterogeneity of network channels, a variety of information services and systems require reliable and uninterrupted operation of telecommunications infrastructure. This problem solved by the introduction of monitoring systems and network management that inform administrators of problem situations (events) that occur due to various errors related to the operation of communications, malicious software or violation of safety rules, etc.

Timely response and elimination of these problems is a part of providing an adequate level of service for end-users and information systems (Grid, Videoconferencing et al.). For such systems to operate properly a guaranteed bandwidth is required.

To manage the Network a set of specialized software management tools deployed in NCC. Protocol ICMP data used in order to control availability of network nodes and evaluation of potential losses in the communication channels; the data obtained by means of open source software tool Nagios. Hardware monitoring, including physical condition of the elements of the kernel and regional nodes in the network, and general evaluation of communication channels utilization is based on the values of SNMP protocol. To collect, store and visualize instrumental data Zabbix system is used. As additional sources of data to evaluate the traffic structure, volume and its distribution between users NetFlow protocol is used. Multiple information systems have to be used for a comprehensive analysis of network status. This is due to the lack of ready open platform that integrates scalable tools to work with most of these protocols enabling the centralized monitoring and processing of data from all network devices.

In the CC FEB RAS the development of a comprehensive system for integrated assessment of the telecommunication network based data monitoring via SNMP and Netflow with the use of open source software [2], [3] is being done. Joint analysis of the data from these two protocols, supplemented by information from the Syslog logs, provides quantitative and qualitative assessment of the Network, both in real time and retrospectively (Fig. 2).

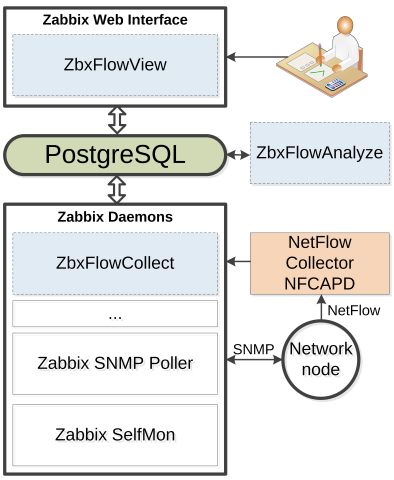

Fig. 2. Structure of information system modules.

3 Scientific Data Storage and Processing Center

In networks with the "star"topology, to which belongs the Regional Computer Network FEB RAS, the data transferred between the segments each time passes through a central hub. If such data streams are permanent (e.g. in distributed information systems, instrumental data collecting, etc.), and the sources of data and their destinations are located in the different network segments or outside, it often produces an excessive load on the main channels associated with duplication of operations. This is especially true for the regions and observation locations, for which limited bandwidth and reliability channels are used (e.g. satellite channels, wireless cellular

communication). An example is the video surveillance system for natural objects, when access to real-time data is required in the same location (Fig. 3).

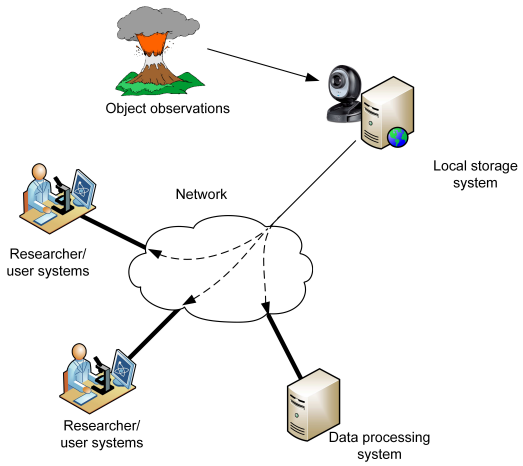

Fig. 3. Example of accessing data at the observation location.

One possible solution to these problems is to migrate the functions of accessing and processing data from the network segments with limited bandwidth channels to data processing centers, located on the network perimeter (DMZ). This ensures an adequate level of reliability of these subsystems and response time of access to data for both internal and external users. In this case, a single connection for data transmission from observation location or regional storage system is established. Multi-user access to data and the results of its processing implemented at the core network that has a comprehensive set of telecommunications and computing hardware and software resources (Fig. 4). This approach allows saving telecommunications resources and providing a guaranteed data collection and delivery.

This approach is widely used in the Regional network of FEB RAS. In particular, it is used for of video surveillance at the volcanoes of Kamchatka and for the network of seismological observations and deformation FEB RAS [4]. The functions of the primary data center in these projects are implemented by the data storage and processing Center of CC FEB RAS (hereinafter - "Data Center"). It also serves as NCC, as well as operation base for hardware and software management systems of a number of centralized and specialized information resources for the benefit of the Far Eastern Branch of the Russian Academy of Sciences and its individual units.

Since 2008, the Data Center has gone through several stages of its technological development [5], and currently consists of the following main components:

1. A data processing system based on IBM BladeCenter H chassis and ten server blades IBM HS23. Each blade contains two processors Xeon E5-2680 8C with a clock speed 2.7GGts, 112 GB of RAM, and network interfaces with the Ten Gigabit Ethernet and Fibre Channel ports.
2. IBM DS3500 Storage System with a total volume of 246.6 TB, consisting of 126 HDD of different capacity and speed - 72 units 3TB each (3.5"SAS 7.2K rpm), 30 units 300GB each (2.5"SAS 15K rpm) 24 units 900 GB each (2.5"SAS 10K rpm). Storage area network (SAN) with a transfer rate 8 Gbps connects server blades with the disk array using Fibre Channel protocol and two switches IBM 44X1920 Brocade.
3. Engineering infrastructure consist of 48 kW Uninterruptible Power Supply APC Symmetra PX, 50 kW diesel generator and air-conditioning systems.

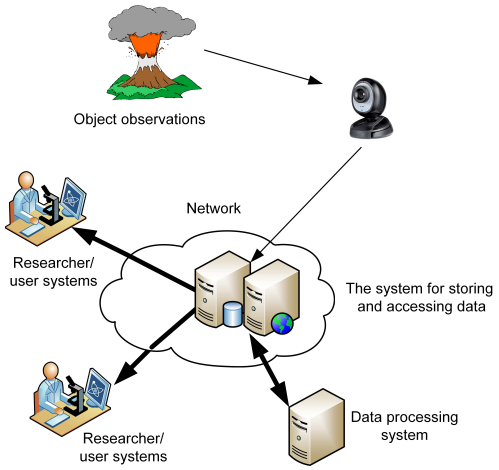

Fig. 4. Example of accessing data using a DMZ.

Data Center is integrated into the Network at a bandwidth of 10 Gbps, which ensures high-speed access to applications and data both for users in the local institutions in Khabarovsk Scientific Center RAS, and for all of the connected FEB RAS institutions. Most of the servers executing the systems and application software are virtual machines created using the virtualization platform VMware vSphere 5.0 Enterprise.

4 Centralized and Specialized Information Systems and Services

The resources of Data Center and regional nodes in the network support the operations of a wide range of information systems. They split into two classes. The first of them, are to form the information environment and the development of technologies for the collective work of scientists, FEB RAS. As an example, the following projects can be mentioned:

- The video conferencing system FEB RAS [6];
- Information system "Grant FEB RAS"designed to support a complex system of competition of scientific projects of the Far Eastern Branch of the Russian Academy of Sciences [7];
- The software platform for creating problem-oriented user interfaces to application software packages for the purpose of calculation in distributed computing systems [8].
- Fault-tolerant information system for cloud storage of scientific data sets [9].
- Instrumental platform for virtual integration of databases [10].

Operations of the systems mentioned above cover most of the institutions and academic staff of the FEB RAS and provides the tools to solve scientific and organizational problems using modern information, telecommunication and computing technologies.

The second class include applied information systems for research in various fields of knowledge.

The most active users of IT-infrastructure organized are the institutions that belong to the Department of Earth Sciences. Primarily they are interested in the use of information technology for the storage and processing of scientific data, operations of instrumental observation networks and conducting a comprehensive study of natural objects and phenomena. The following examples of such established systems are:

1. **Information System "Remote Monitoring of Active Volcanoes of Kamchatka and the Kuril Islands"(VolSatView).** To reduce the risk of collision of aircraft with ash clouds in different parts of the world where there are active volcanoes, volcano observatories designed to monitor these volcanoes. In 1993 Kamchatkan Volcanic Eruption Response Team (KVERT) was created. Since 2010 it is a part of the Institute of Volcanology and Seismology (IVS FEB RAS), and serves as the Volcano Observatory of the Russian Federation for the international air navigation community information on volcanic activity in the Far East [11], [12]. Since 2011, specialists IVS FEB RAS, IKI, CC FEB RAS, FGBU SIC "Planet"are working on the creation of a special information system that provides for the collection, storage and processing of satellite data for volcanic activity monitoring in the region of Kamchatka and the Kuril Islands. The system is called Satellite service "Monitoring of active volcanoes of Kamchatka and the Kuril Islands"(VolSatView). In addition, the system provides a variety of tools for the analysis of this information, joint analysis using data sets of instrumental observations of ground networks including data obtained from third-party information systems (Fig. 5) [13].

   The information system is distributed. Physically, its nodes with installed software are located in three Russian cities: Moscow, Khabarovsk and Petropavlovsk-Kamchatsky. The reason for this is the presence of already functioning data centers and storage, satellite receiving stations and users of the data. Virtual network organized to exchange data and work with the system components using dedicated channels of FEB RAS Network and FGBU SIC "Planeta". The main network activity falls into segments DC SIC Planet - CC FEB RAS - IKI, in which circulates the flow of data associated with the completion of satellite data archives. Ability to install QoS at network channels, provided an opportunity to receive data and work with the system for groups of researchers in Kamchatka, where there is a limitation to the resources associated with the use of satellite communication channels.

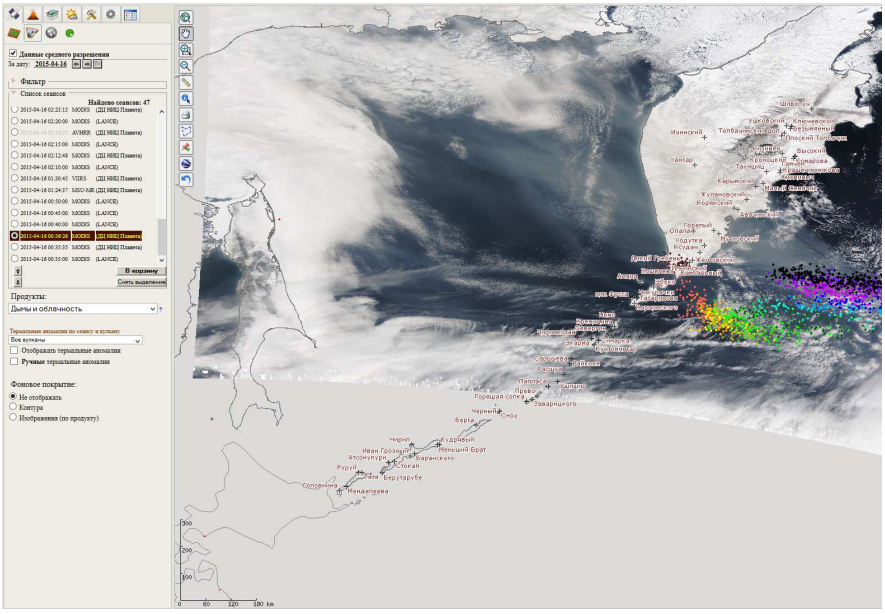

Fig. 5. Information System "VolSatView": Example screen form.

2. **Automated Information System "Signal".** In 2009, FEB RAS approved the Comprehensive program of basic scientific research "Modern geodynamics, active geostructure and natural hazards Russian Far East."Its main objective was the development and application of modern methods of research the movements of crustal blocks, deep structures, etc. based on the monitoring of deformations and seismicity of different scales at the junction of the North American, Pacific, Okhotsk and Amur lithospheric plates. As a part of the planned activities in the Far East a co-operation of ITG FEB RAS, IMGG FEB RAS, IAM RAS, IGNM FEB RAS and CC FEB RAS started to create a network of continuous seismological [16] and deformation observations [17] In addition to the means of recording (seismic, geodetic GPS / GLONASS receiver, etc.) the observation network included local data archives, set at observation locations and a main Data center (in CC FEB RAS). Transmission of data from the field observations to the Data center is performed using network resources and networks of mobile service providers.
An automated information system "Signal"was developed to manage the observation network infrastructure and working with instrumental data (Fig. 6) [18],[19]. Its architecture allows integrating additional components for specialized processing of scientific information into the system.
At present, the functionality of the system includes software modules to work with sets of seismic and geodynamic data, archives of surveillance of volcanoes of Kamchatka, including computing components for modeling the motion path of eruption clouds and plumes and study geodynamic phenomena and processes based on GPS / GLONASS observations.

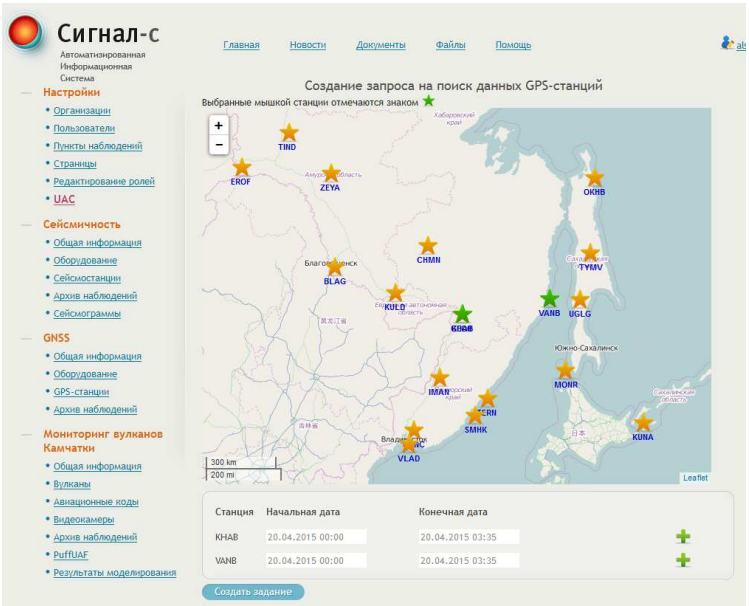

Fig. 6. Information System "Signal": Example screen form

5 Conclusion

Over the past 10 years, the formation of basic telecommunications and computing infrastructure completed in the Far Eastern branch of Russian Academy of Sciences. It provides an opportunity to scientists to conduct research using the most advanced information technologies.

As a result, it allowed the implementation of projects related to the construction of distributed information systems for large-scale research in the field of Earth Sciences [13], [20], [21].

The need to solve the problems of data collection, storage and processing of information required to deploy core data center, which is developing technologies for the integration of scientific data [8], their distributed processing and storage, including the use of cloud computing [9]. Concentration of these data, the formation of modern information services based on common standards and technologies allows us to consider integration with external systems to solve current scientific and organizational tasks. Further development of telecommunication systems and increase the capacity of the network channels will allow in the future considering the formation of a common scientific information environment in the Far East of Russia.

Acknowledgements. This work was partially supported by grants FEB RAS No.15-I-4-071, RFBR No.13-07-12180, 15-37-20269 mol_a_ved and Far Eastern Federal University No.14-08-01-05_m.

References

1. Smagin S.I., Sorokin A.A., Makogonov S.V.: Corporate network FEB RAS: New ways of development. In: Proceedings of the IASTED International Conference on Automation, Control, and Information Technology - Information and Communication Technology, ACIT-ICT. July 15-18. 2010 pp. 202-205. Novosibirsk.(2010)
2. Sorokin A.A., Korolev S.P., Tarasov A.G., Verkhoturov A.L.: Research and development of information systems for monitoring of telecommunication networks based on open standards and technologies. In: Proceedings of the XIX Russian conference "Distributed Information and Calculation Resources"(DICR'2013) electronic materials. pp.213-218. Novosibirsk: ICT SB RAS, 2013. - Vol. I. ISBN 978-5-905569-0. (2013)
3. Sorokin A.A., Tarasov A.G., Korolev S.P.: Automated information system for comprehensive monitoring of telecommunication network. Systems and measures in Informatics
4. Khanchuk A.I., Konovalov A.V., Sorokin A.A., Korolev S.P., Gavrilov A.V., Bormotov V.A., Serov M.A.: Instrumental and of information-technological provision of seismological observations in the Russian Far East. Herald of FEB RAS,No.3. 2011. pp. 127-137.(2011)
5. Makogonov S.V.: Khabarovsk Data Center Project (preprint). - Khabarovsk: CC FEB RAS, p. 30. (2008)
6. Makogonov S.V., Pogorelov S.A.: The development of video conferencing systems on the example of Videoconferencing system of FEB RAS. In: Proceedings of Science practical conference Information technologies and high-performance computing. Khabarovsk: KHSTU, pp. 211-216. (2013)
7. Smagin S.I., Michailov K.V., Krivosheev A.I., Savin S.Z.: On the technology to create interactive information systems Informatics and Control Systems No.1, 2011, pp. 115-120. (2011)
8. Tarasov A.G., Mikhailov K.V., Sorokin A.A., Korolev S.P.: Technology to create and maintain task-oriented interfaces for distributed computing systems. Informatics and control systems. No. 39, 2014, pp. 70-79 (2014)
9. Sorokin A.A., Korolev S.P., Smagin S.I., Polyakov A.N.: Layout of fault-tolerant information system for cloud storage of scientific data sets. Computational technologies. . vol. 18. No 1. 2013. pp. 87-95 (2013)
10. Polyakov A.N., Poida A.N., Sorokin A.A., Smagin S.I., Korolev S.P.: Development of programs for virtual integration of distributed data sources for creating professionally-oriented large-scale information environment. Informatics and Control Systems. No.3, 2013, pp. 152-160. (2013)
11. Girina O.A.: 15 years of activity of the Kamchatka Volcanic Eruption Response Team (KVERT) In: Proceedings of the Conference on Volcanologist Day, March 27-29, 2008. Petropavlovsk-Kamchatsky: IVS SB RAS, pp. 52-59. (2008)
12. Girina O.A., Gordeev E.I.: The project KVERT - reduction of volcanic hazards to aviation in explosive eruptions of volcanoes in Kamchatka and Northern Kuriles. Herald of FEB RAS. No.2, 2007, pp. 100-109.(2007)
13. Efremov V.Y., Girina O.A., Kramareva L.S., Lupyan E.A., Manevich A.G., Melnikov D.V., Matveev A.M., Proshin A.A., Sorokin A.A., Flitman E.V.: Creating an information service "Remote monitoring of volcanic activity in Kamchatka and the Kuril Islands". Modern problems of remote sensing of the Earth from space.No.5, 2012. pp. 155-170.(2012)
14. Smagin S.I., Lupyan E.A., Sorokin A.A., Butsev M.A., Verkhoturov A.L., Girina O.A., Efremov V.Y., Kramareva L.S., Proshin A.A., Tolpin V.A.: Information system for processing data of satellite observations of Russia Far East for science investigations in different areas of knowledge. Modern problems of remote sensing of the Earth from space. No.1, 2013, pp. 277-291. (2013)

15. Shokin Y.I., Antonov V.N., Dobretsov N.N., Kikhtenko V.A., Lagutin A.A., Smirnov V.V., Chubarov D.L., Chubarov L.B.: A distributed system for receiving and processing satellite data of Siberia and the Far East. Current state and prospects of development. Modern problems of remote sensing of the Earth from space. No.5, 2012 pp. 45-54. (2012)
16. Khanchuk A.I., Konovalov A.V., Sorokin A.A., Korolev S.P., Gavrilov A.V., Bormotov V.A., Serov M.A.:. Instrumental and of information-technological provision of seismological observations in the Russian Far East. Herald of FEB RAS. No.3, 2011. pp. 127-137. (2011)
17. Bykov V.G., Bormotov V.A., Kokovkin A.A., Vasilenko N.F., Prytkov A.S., Gerasimenko M.D., Shestakov N.V., et al.: Start building a unified network of geodynamic observations, FEB RAS. Herald of FEB RAS.No.4, 2009. pp. 83-93.(2009)
18. Sorokin A.A., Korolev S.P., Mikhaylov K.V., Konovalov A.V.: "SIGNAL-S Automated information system for seismological data processing. In: Proceedings of the 1st Russia and Pacific conference on computer technology and applications, RPC 2010. September 06-09. 2010. pp. 283-284. Vladivostok: Institute of Electrical and Electronics Engineers (IEEE). (2010)
19. Korolev S.P., Sorokin A.A., Verkhoturov A.L., Konovalov A.V., Shestakov N.V.: Automated information system to work with instrumental data of the regional network of seismological observations FEB RASuments. No.3, Vol. 50,2014. pp. 30-41.(2014)
20. Shestakov N. V., Ohzono M., Takahashi H., Gerasimenko M. D., Bykov V. G., Gordeev E. I., Chebrov V. N., Titkov N. N., Serovetnikov S. S., Vasilenko N. F., Prytkov A. S., Sorokin A. A., Serov M. A., Kondratyev M. N., Pupatenko V. V.: Modeling of coseismic crustal movements initiated by the May 24, 2013, M w = 8.3 Okhotsk deep focus earthquake. Doklady Earth Sciences. vol. 457. No.2. 2014. pp. 976-981. (2014)
21. Khanchuk A.I., Safonov D.A., Radziminovich Ya.B., Kovalenko N. S., Konovalov A.V., Shestakov N.V., Bykov V.G., Serov M.A., Sorokin A.A.: The largest recent earthquake in the Upper Amur region on October 14, 2011: First results of multidisciplinary study. Doklady Earth Sciences. vol. 445. No.1. 2012. pp. 916-919.(2012)

Архитектура Программного Комплекса Интеллектуального Облачного Сервиса Мониторинга Состояния и Управления для Удаленных Распределенных Объектов

Д.В. Волков[1] и А.М. Епихин[1]

[1] СКТБ "Наука"КНЦ СО РАН, Красноярск
vdv@kis.krasnoyarsk,AndrewE@dmdevelopment.ru

Аннотация. В докладе рассматриваются вопросы создания программного комплекса интеллектуального облачного сервиса, реализующего новый подход к построению автоматизированных систем для мониторинга и управления распределенными удаленными объектами различного назначения. Представлены принципы построения, функциональная спецификация, архитектура программного комплекса. Предложены подходы к обеспечению надежности и производительности. Приведены примеры прикладных систем, реализованных на платформе облачного сервиса.

Ключевые слова: Облачный сервис, распределенные объекты, удаленный мониторинг, программное обеспечение.

1 Актуальность

Массовое применение взаимодействующих через интернет автоматических или автоматизированных устройств - общемировой тренд. Хорошие перспективы имеет интеллектуальный облачный сервис мониторинга состояния и управления удаленных распределённых объектов различной природы [1]. Это могут быть природные и технические объекты, физические лица с носимыми датчиками контроля физиологических параметров и проч.

По сути речь идет о новой парадигме создания информационных систем, характеризующихся низким уровнем затрат на разработку и эксплуатацию, простотой использования, богатыми возможностями настройки и развитыми интеллектуальными сервисами эксплуатации. Для этого необходимо иметь возможность просто, быстро, с минимальными затратами развернуть для распределенных в пространстве объектов и субъектов (пользователей и специалистов, участвующих в эксплуатации объектов) интеллектуальный программно-аппаратный комплекс (ПАК), позволяющий обеспечить информационную поддержку всех участников процессов взаимодействия с объектом.

В состав программно-аппаратного комплекса входят: объекты управления (ОУ), облачный сервис, служба технической поддержки сервиса, служба сопровождения прикладных систем (отраслевых решений), субъекты управления - администратор ПАК, оператор управления - сотрудник, имеющий автоматизированное рабочее место (АРМ), АРМы организаций, выполняющих функции обслуживания объектов (рис. 1).

Основой облачного сервиса является программный комплекс, который позволяет: организовать каналы передачи данных для взаимодействия объектов управления, операторов и внешних систем, принимать, хранить и обрабатывать данные, поступающие от объектов управления, операторов и внешних систем, предоставлять операторам сервисы по организации удаленного мониторинга и управления объектами.

На объектах, подлежащих контролю, устанавливаются контроллеры с комплексом датчиков. Датчики опрашиваются контроллером, а построенная на его основе система управления формирует команды исполнительным устройствам объекта. Объект имеет канал связи

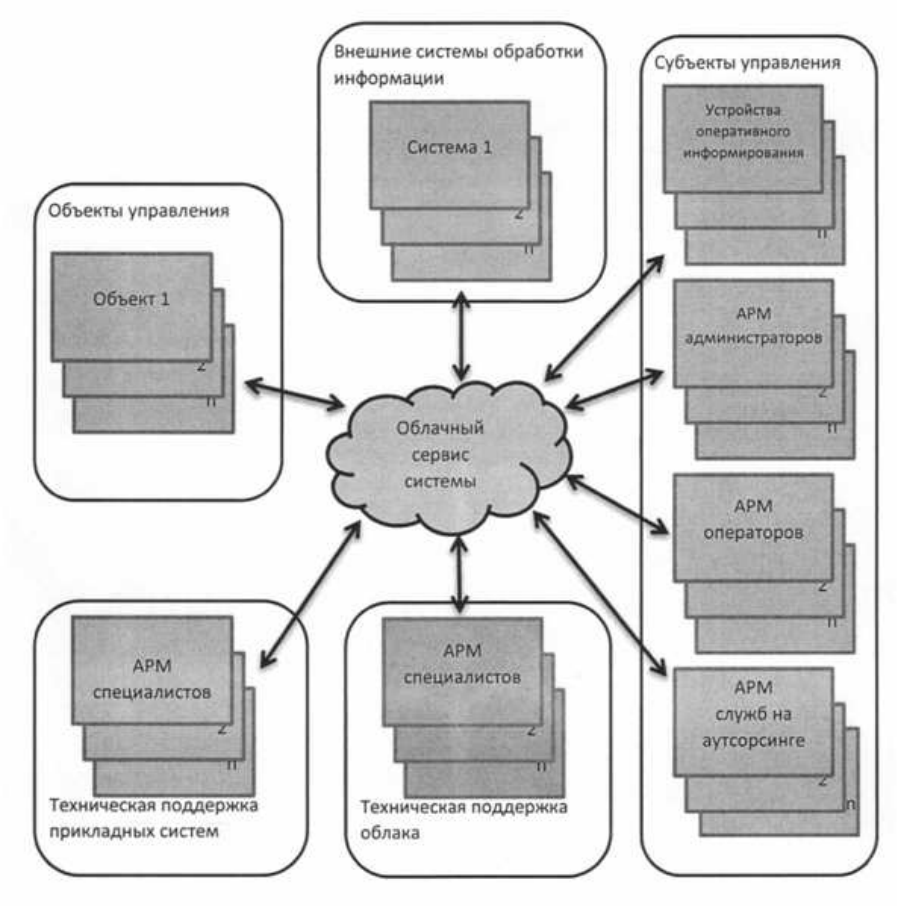

Рис. 1. Схема взаимодействия компонент ПАК облачного сервиса.

с облачным хранилищем данных, в котором накапливаются данные для обработки на более высоком уровне. Целями обработки могут быть: анализ данных, выработка управляющих воздействий, интеллектуальные процедуры принятия решений, организация меж объектного взаимодействия, формирование новых логик мониторинга и управления. Предполагается поддержка автоматических и автоматизированных систем с операторами процессов мониторинга и управления.

2 Характеристика объектов автоматизации

Определим объект, как квазистационарную организационно-техническую систему, состояние которой в каждый момент времени характеризуется набором значений параметров (вектором состояния). Термин организационно-техническая означает, что обязательным элементом системы, обеспечивающим ее функционирование, является человек - оператор. Термин квазистационарная означает, что объект может быть как недвижимым в пространстве, так и подвижным, но существенно, что параметры траектории его движения не являются целями управления.

Объект, для которого необходимо решить задачу контроля параметров в процессе функционирования, определим как объект мониторинга. Объект, для которого необходимо решить задачу управления параметрами в процессе функционирования, определим как объект

управления. Объект, для которого необходимо одновременно решить задачи контроля параметров и управления параметрами в процессе функционирования, определим как объект мониторинга и управления. Текущее состояние объекта описывается вектором состояния, а его изменение во времени описывается траекторий в пространстве состояний.

Объект и его подсистемы характеризуются ассоциированными с ними наборами параметров. Каждый параметр принимает значения из множества возможных для заданного параметра значений. Часть параметров может определяться как уставки (целевые состояния). Для целей построения алгоритмов управления можно программировать логические высказывания, используя параметры в качестве переменных. Момент принятия высказыванием значения "истина"или "ложь"определим, как момент наступления события, ассоциированного с этим высказыванием.

Параметры объекта, которые измеряются с помощью датчиков, определим как наблюдаемые. Элемент объекта, который может изменять свое состояние под воздействием управляющей команды, определим как исполнительный механизм. Возможны различные типы исполнительных механизмов.

Стационарные объекты имеют измеряемые и/или управляемые параметры, микропроцессорную систему, позволяющую управлять датчиками/или исполнительным оборудованием. Могут работать в автономном, автоматическом или полуавтоматическом (с локальным оператором) режимах. При наличии онлайн соединения могут взаимодействовать с внешней диспетчерской (центром управления) или другими внешними системами обработки информации. Существует возможность периодически или на постоянной основе организовать канал связи удаленного объекта или субъекта с облачным сервисом через интернет.

3 Функции программного комплекса

Облачный сервис предназначен для организации взаимодействия объектов, субъектов, внешних систем и вспомогательных служб в процессе решения задач автоматизированного удаленного мониторинга и управления объектами.

В облачной среде размещается специализированное программное обеспечение, которое позволяет решать следующие задачи:

1. Организовать каналы передачи данных для взаимодействия с объектами управления, субъектами управления и внешними системами.
2. Настроить (сконфигурировать) Систему.
3. Принимать, хранить и обрабатывать данные, поступающие от объектов управления и внешних систем в соответствии с заданной логикой обработки.
4. Предоставлять субъектам управления сервисы по организации удаленного мониторинга и управления объектами посредством специализированых АРМ.
5. Предоставлять субъектам управления и сотрудникам Службы сопровождения сервисы по взаимодействию (контрактинг, биллинг, техподдержка и проч.).

Для решения указанных задач специализированное облачное ПО предоставляет зарегистрированным субъектам управления функциональные сервисы через специализированное ПО поддержки клиентского рабочего места - консоль администрирования.

Консоль администрирования позволяет выполнить следующие функции:

1. Зарегистрировать/редактировать параметры нового ОУ.
2. Настроить (конфигурировать) систему управления ОУ (задать параметры, логику функционирования, условия применения, правила распознавания и обработки нештатных ситуаций и др.).

3. Зарегистрировать пользователей и настроить (сконфигурировать) рабочие среды пользователей.
4. Настроить среду и правила взаимодействия объекта и субъекта.
5. Зарегистрировать специалистов Служб сопровождения и правила их взаимодействия с Системой.

АРМ пользователя позволяет выполнить следующие группы функции:

1. Вносить разрешенные изменения в конфигурацию Системы (настройки пользователя):
   (a) задать/изменить закон управления,
   (b) задать/изменить правила взаимодействия объекта и субъекта,
   (c) ввести в рабочую среду пользователя с ограниченными правами, задать/изменить его функциональные возможности (только смотреть, смотреть и реагировать, устанавливать правила реагирования и проч.).
2. Осуществлять визуальный, звуковой контроль на виртуальных мониторах.
3. Распознавать штатные или нештатные ситуации, требующие вмешательства, и формировать соответствующие им команды управления.
4. Формировать отчеты о процессах мониторинга и управления.
5. Осуществлять взаимодействие со Службами сопровождения облачного сервиса и прикладного решения (зарегистрировать инцидент, сформировать запрос на его обработку, решить вопросы по договору и оплате и проч.).

Предполагается следующий порядок работы:

1. Регистрация пользователя в сервисе (заведение аккаунта).
2. Установка настройка ПАК на объекте.
3. Подключение (настройка каналов обмена данными).
4. Конфигурирование системы управления.
5. Штатная эксплуатация системы управления пользователем (группой пользователей).
6. Периодическое взаимодействие пользователя со Службой сопровождения по вопросам техподдержки, контрактинга, оплаты и проч.

4 Архитектура программного комплекса

Программный комплекс состоит из четырех основных подсистем:

1. Подсистема сбора данных об объектах мониторинга (состоит из агентов и сервисов сбора данных).
2. Подсистемы визуализации данных и мониторинга.
3. Хранилищ данных и подсистемы анализа.
4. Ядра системы по управлению метаданными об объектах мониторинга.

Кроме того присутствует инфраструктура по управлению агентами сбора данных и клиентами визуализации.

Основная идея системы состоит в том, чтобы отойти от визуализации и анализа первичных данных, построить модель анализа, наделённую семантикой, адекватной предметной области объекта мониторинга, и уже отображение и мониторинг данных выполнять в соответствии с этой моделью. Настройка модели становится возможной путём конфигурирования объектов мониторинга, а в процессе работы системы осуществляются преобразования исходных данных, полученных агентами в такие показатели мониторинга, которые были

запрошены пользователем в процессе конфигурирования комплекса его персональных настроек.

Вся система построена с использованием принципов CQRS [2] и асинхронного взаимодействия. Программный комплекс реализован в виде множества сервисов, развёрнутых на различных узлах (серверах, рабочих станциях, специализированных контроллерах).

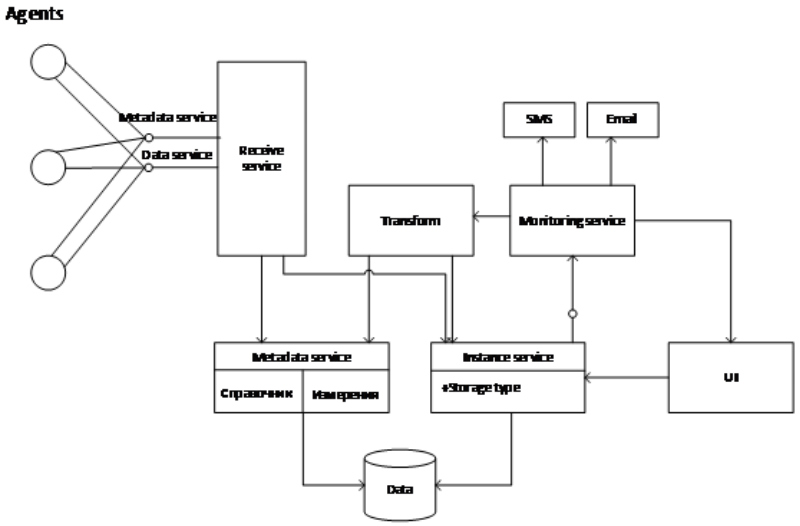

Рис. 2.

Ниже перечислены все компоненты системы.

Сервисы\Агенты сбора данных (DataAgent) - служат для регистрации и сбора данных с объектов мониторинга. В системе мониторинга предполагается использование множества специализированных агентов как программных, так и аппаратно- программных. Например, могут быть созданы агенты для получения регистрации параметров объектов с использованием различных датчиков и контроллеров, агенты для сбора данных OPC-сервера, получения событий операционной системы или сетевого оборудования.

При построении агентов используются следующие идеи:

1. Агент может возвращать данные различного типа, вида и структуры от скалярных значений до мультимедиапотоков с использованием различных контейнеров и методов кодирования.
2. Агент сам описывает структуру и типы данных, которые возвращает, а также допустимые операции.
3. Агент имеет специальный API для настройки и конфигурирования.
4. Агент может быть реализован как в виде активной, так и пассивной стороны взаимодействия (push и pull реализация обмена данными).

Метаданные, которые возвращает агент о собираемых им данных, далее используются ядром системы для построения модели анализа и преобразования для получения нужных показателей.

Сервис получения данных служит для получения данных от агентов сбора данных.

Модуль управления метаданными служит для хранения метаданных (как полученных от агентов, так и внесённых в систему в процессе конфигурирования), управлению ими, построения модели анализа.

Модуль преобразования данных служит для преобразования, агрегирования данных и сохранения их в разные виды хранилищ.

Подсистема визуализации состоит из Web интерфейса для отображения и интерактивного анализа данных и некоторых сервисов, обеспечивающих высокую доступность некоторых видов данных (потоковое видео, сторонние файлы, мультимедиа информация).

Конфигурирование и пользовательский ввод данных - web-интерфейс для ввода метаданных и некоторого вида справочной информации.

Сервисы мониторинга и оповещения служат для мониторинга за событиями и отклонениями как исходных, так и в преобразованных\агрегированных данных. Например, с помощью этих сервисов можно настроить получение email, sms при возникновении критического события на объекте мониторинга или отклонении статистического показателя от заданной нормы.

Хранилища данных служат для хранения как исходных данных, так и данных после преобразования. В системе возможно использование различных типов хранилищ, как не реляционных (некоторые виды исходных данных), реляционных, и многомерных (некоторые виды агрегированных показателей).

5 Примеры прикладных систем на основе сервиса

У предлагаемого сервиса просматривается множество возможных применений. Для примера рассмотрим три варианта: система диспетчеризации тепловых пунктов многоквартирных домов, обслуживаемых управляющей компанией, автоматизированная система диагностики сложного оборудования в процессе эксплуатации, система мониторинга состояния сложных инженерных сооружений.

Система диспетчеризации тепловых пунктов многоквартирных домов, обслуживаемых управляющей компанией. Поддерживает сервисы контроля доступа к объекту, съема показателей с приборов учета тепловой энергии, контроля параметров работы автоматизированных систем, обеспечивающих регулирование отбора тепла от теплоносителя на уровне возмещения теплопотерь дома, контроль других важных параметров теплового контура дома и проч.

Автоматизированная система диагностики сложного оборудования в процессе эксплуатации. Система позволяет обеспечивать контроль нахождения параметров оборудования в установленных пределах в процессе функционирования, распознавать и обрабатывать нештатные ситуации. Информировать специалистов о возникновение нештатных ситуаций, требующих вмешательства (остановки оборудования, обслуживания или перевода в другой безопасный режим функционирования). Информирование о нештатной ситуации сервисной службы позволяет своевременно предпринять необходимые действия и предотвратить потери, связанные с неуправляемым развитием аварийной ситуации и выходом оборудования из строя.

Система мониторинга состояния сложных инженерных сооружений: мостов, эстакад, гидротехнических сооружений и проч. Накопление информации, распознавание нештатных ситуаций, информирование обслуживающего персонала, диспетчерские службы.

6 Перспективы развития

В настоящее время создана первая версия ПО облачного сервиса, идет создание и апробация ряда прикладных информационных систем на его основе. По результатам апробации и

опытной эксплуатации будет подготовлена спецификация на рабочую версию программного комплекса. Прорабатываются вопросы обеспечения информационной безопасности, как критически важной характеристики для программного обеспечения систем рассматриваемого класса [3].

Список литературы

1. Бурлаков А.В., Волков Д.В., Епихин А.М., Интеллектуальный облачный сервис мониторинга состояния и управления для удаленных распределенных объектов // Фундаментальная информатика, информационные технологии и системы управления: реалии и перспективы. FIITM-2014: материалы международной научно-практической конференции Красноярск: Сибирский федеральный университет, 2014. – С. 61-68.
2. URL: `http://martinfowler.com/bliki/CQRS.html` (дата обращения: 18.05.2015).
3. Марк Браун, Пять шагов к обеспечению безопасности встраиваемых систем// Современные технологии автоматизации 2013, – №3, – с. 14-22.

Применение ГИС и WEB-Технологий для Создания Геоинформационной Системы «Инвестор»

А.С. Гаченко., А.А. Михайлов., Г.М. Ружников., А.Е. Хмельнов.

Федеральное государственное бюджетное учреждение науки Институт динамики систем и теории управления имени В.М. Матросова СО РАН, Иркутск, Россия
gachenko@icc.ru

Аннотация. Для повышения инвестиционной привлекательности территорий муниципалитетов и увеличения бюджета городов в России активно разрабатываются и внедряются инвестиционные карты или геоинформационные системы (ГИС) «Инвестора». Создание таких систем требует интеграции информации о городской территории, содержащейся в базах данных (БД) и на электронных картах существующих муниципальных информационных систем (МИС). Геоинформационные системы «Инвестора» являются уникальными для каждого города и не могут тиражироваться. Поскольку содержащие исходные данные МИС являются специфичными для каждого города, а их структура определяется сложившимися особенностями муниципального управления и длительной историей разработки. В данной работе рассматривается опыт создания ГИС «Инвестора» г. Иркутска, на основе оригинальных технологий и технических решений.
Цель разработки ГИС «Инвестора» г. Иркутска – создание для потенциальных инвесторов благоприятных условий, путем обеспечения открытости и доступности информации, необходимой для выбора на территории города объектов представляющих возможный интерес.

Ключевые слова: ГИС «Инвестора», инвестиции, муниципальные информационные системы, привлечение средств инвестора, информационно обеспечение, ГИС, городское хозяйство, публикация данных в интернет.

1 Введение

Для повышения инвестиционной привлекательности территорий городов на протяжении ряда последних лет активно разрабатываются и внедряются муниципальные геоинформационные системы (ГИС) «Инвестор» [3,4,5,6]. Инвестиции являются важной составляющей успешного развития городской экономики. Консалтинговые и оценочные компании, девелоперы, предприниматели используют эти ГИС для получения информации об инвестиционных возможностях территорий, повышения скорости превращения инвестиционной идеи в работающий бизнес. Данные ГИС «Инвестор», как правило, уникальны, поскольку содержат специфичные для каждого города пространственные и тематические данные, а их структура и функциональность определяется особенностями территориального управления.

2 Структура ГИС «Инвестор»

В городе Иркутск ГИС «Инвестор» интегрирует актуальные общедоступные пространственные и тематические данные об объектах территориальной инфраструктуры города формируемых профильными комитетами и подведомственными им организациями, участвующими в инвестиционном процессе [2,3,4].

Структурно ГИС «Инвестор» включает: Web-браузер, модуль отображения данных (Web-сервер под управлением Internet Information Services, библиотека Leaflet), модуль администрирования (АИС подготовки данных), конвертор из различных форматов в формат SMD, WMS сервисы (космоснимки сторонних организаций), хранилище пространственных

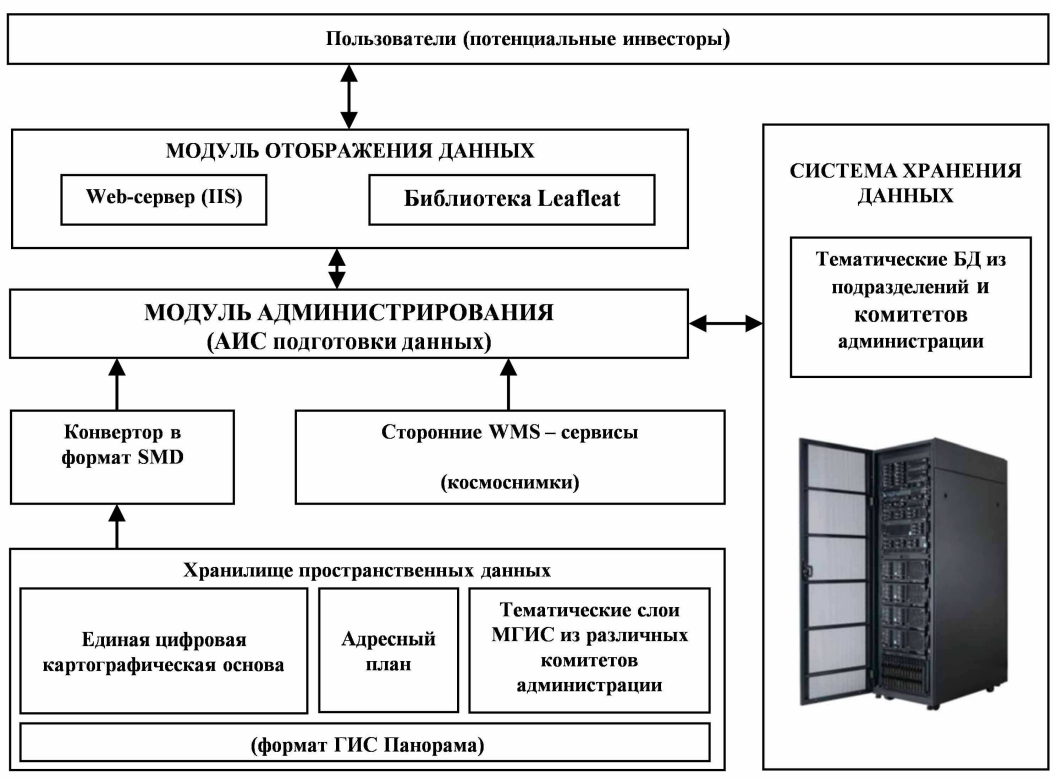

Рис. 1. . Структура ГИС «Инвестор»

данных (ЕЦКО, адресный план, тематические слои подразделений), система хранения (тематические БД подразделений и ГИС «Инвестор»).

Функционально ГИС «Инвестор» обеспечивает визуализацию всех доступных для отображения слоёв существующей муниципальной геоинформационной системы (МГИС) города Иркутска, с возможностью их включения/отключения, выбора и поиска объекта или группы объектов, отображения семантической информации, интегрированной по всем включённым в выборку слоям карты.

Основным элементом ГИС «Инвестор» является модуль отображения данных, который разработан с использованием технологии декларативных спецификаций. Остальные элементы структуры системы предназначены для администрирования и организации взаимодействия с тематическими муниципальными информационными системами и БД (Рис. 1).

В качестве базовой топоосновы используется цифровой «Адресный план» г. Иркутска, который регулярно обновляется на основе «Дежурного цифрового топографического плана» города [4]. Информация для ГИС «Инвестор» автоматически формируется из тематических слоёв муниципальной геоинформационной системы г. Иркутска (МГИС). Кроме того, у системы существует собственная БД, которая предназначена для хранения служебной информации, а также для быстрого добавления табличных данных, получаемых из подразделений администрации г. Иркутска и других внешних источников.

Большинство цифровых тематических карт, используемых в г. Иркутске, представлены в формате ГИС «КАРТА», которые для ускорения обработки Web-модулем конвертируются в формат SMD.

ГИС «Инвестор» использует следующие цифровые тематические карты (слои):

- Адресный план (официальная утвержденная выверенная карта города).
- Кадастровое деление.
- Правила землепользования и застройки.
- Проекты планировок.
- Транспортный слой.
- Слой расположение торговых организаций.

Разработан удобный интерфейс управления узлами и слоями, то есть можно управлять не только каждым слоем, но и их группами, устанавливать степень прозрачности слоев на карте и менять очередность их отображения.

Состав отображаемой табличной информации ГИС «Инвестор»:

Таблица 1. Данные ГИС «Инвестор»..

Коммерческие объекты	Гостиницы Магазины Торговые центры
Население	Агрегированные данные по жителям г. Иркутска
Социальные объекты	Дополнительное образование Дошкольное образование Объекты здравоохранения Культурные объекты Спортивные объекты Муниципальные образовательные учреждения
Торги муниципального имущества	Объекты муниципальной недвижимости выставленные на торги
Транспорт	Автобусные остановки

При выборе на интерактивной карте конкретного объекта недвижимости автоматически отображается информация о нём.

Формирование и наполнение хранилища пространственных данных осуществляется с использованием геоокодирования. Также существует возможность обновления данных в режиме добавления на карту новых слоев с уже существующей в них семантической информацией.

3 Модуль отображения ГИС данных

Пользовательский интерфейс представления прост и визуально компактно организован. Основным элементом интерфейса является цифровой «Адресный план» города. Для отображения цифровых карт в ГИС «Инвестор» используется библиотека Leaflet [7], которая представляет собой JavaScript библиотеку для публикации карт в мобильных и настольных браузерах. Leaflet предоставляет возможность разделения на модули и применения различных расширений. Клиентская часть ГИС «Инвестор» включает в себя:

- Модуль работы с картой

- Модуль поиска
- Модуль управления слоями
- Модуль выбор источника данных
- Модуль отображения слоев

Модуль поиска позволяет получать координаты о географическом объекте по его адресу и выставить позиционирование карты по центру найденного объекта. В данном модуле реализована поддержка API геокодирования Web-сервисов:

- Карты Google
- Карты Яндекс
- 2 Гис

Пользователь может отметить на карте любой объект и получить по нему в отдельном окне имеющуюся о нём информацию из Адресного плана, а так же и из публичных источников: кадастровой карты, 2 Гис.

В ГИС «Инвестор» для публикации картографических данных на стороне клиента используется формат представления векторных данных SMD (Static Map Data). Данный формат позволяет эффективно (по времени отображения и загрузки информации в память) визуализировать пространственные объекты. Перед публикацией данных подсистема «Администрирования» автоматически проводит конвертацию картографических материалов именно в SMD.

Панель управления пользовательского интерфейса содержит элементы управления ГИС функциональностью (навигация, масштабирование, измерение расстояния между объектами), выбора объектов карты, в том числе по произвольному контуру, отображения легенды слоев карты, отображения семантической информации в области отчётов, поиска объектов на карте по адресу и другим атрибутам. Пользователи в ГИС «Инвестор» имеют возможность дополнительно просматривать фрагменты на космоснимках. Разработана технология мониторинга данных на картографических сервисах, которая позволяет просматривать выбранный фрагмент карты на сторонних картографических онлайновых сервисах:

- Карты Google
- Карты Яндекс
- OpenStreetMap (OSM) – свободная, бесплатная вики-карта мира
- Публичная кадастровая карта

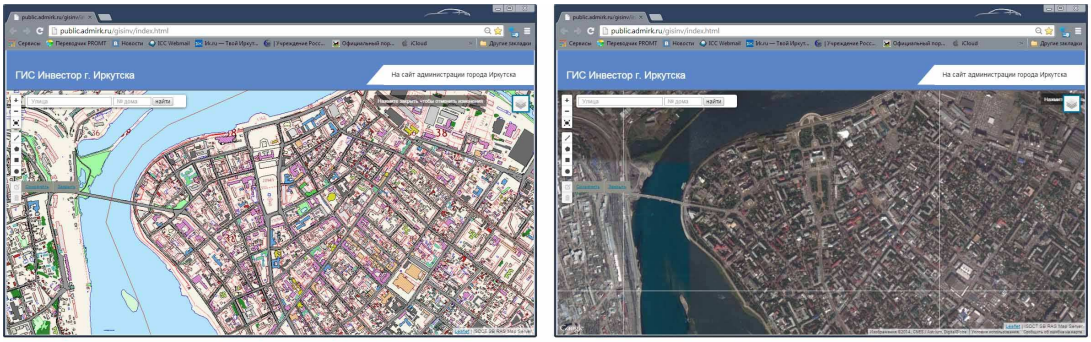

Рис. 2. Отображения выбранного фрагмента на космоснимках Google (слева оригинал, справа картографический сервис Google)

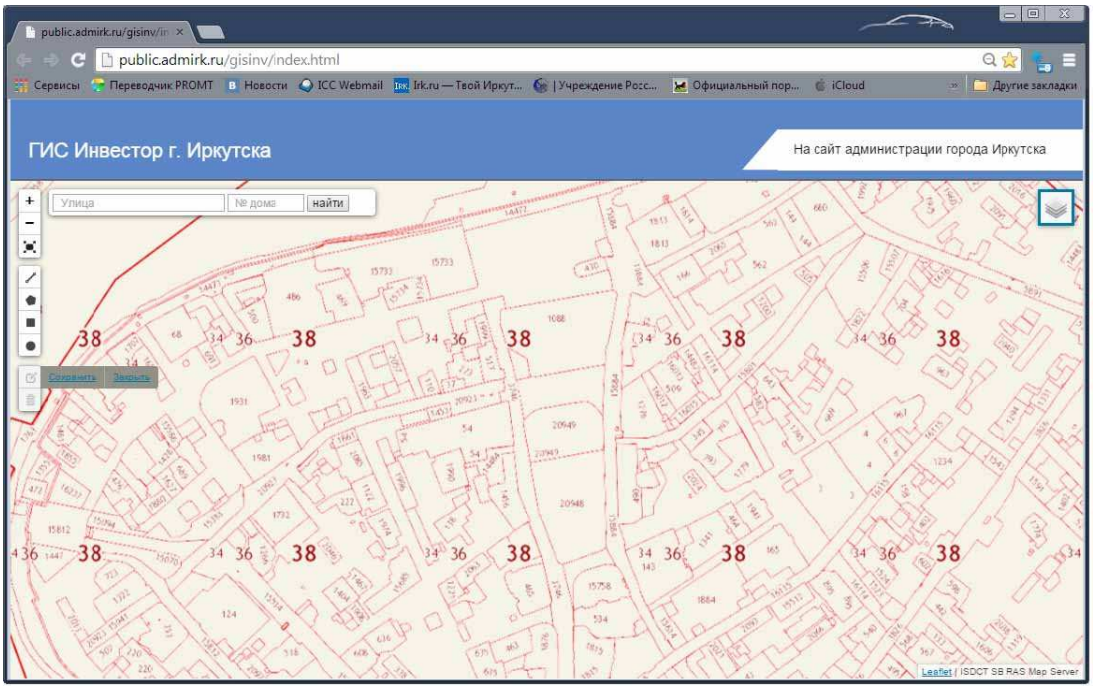

Рис. 3. Отображения выбранного фрагмента на публичной кадастровой карте Росреестра

Наиболее ресурсоёмким и информативным инструментом, реализованным в интерфейсе пользователя, является блок построения запросов, который позволяет создавать отчетные выборки. Для вывода информации об объектах из базы данных используется технология декларативных спецификаций [2]. Одной из разновидностей запросной системы является выборка по произвольной (ограничивающий многоугольник) области (Рис. 4). При выполнении такого запроса формируется отчёт по объектам карты указанной области из отмеченных в дереве информационных слоёв. По каждому из выбранных слоёв вычисляются итоговые значения, а также может отображаться список объектов данного слоя, если выборка выполняется по ограниченной области, содержащей не слишком много таких объектов. При выводе информации об объектах слоя в отчёт включаются карточки записей из связанной со слоем таблицы БД. Для формирования таких карточек также используются декларативные спецификации структуры БД [1,2].

В целях исключения длительного формирования отчетных данных, в связи с выбором большого количества объектов, разработан блок прогнозирования времени формирования отчетов и оптимизации запросов с предоставлением пользователю информации о количестве объектов включенных в его запрос.

4 Модуль администрирования ГИС «Инвестор»

Модуль администрирования ГИС «Инвестор» предназначен для конфигурирования модуля отображения данных и подготовки информации для него из различных тематических и графических баз данных. Для этого применяется инструментальная среда «ГеоАРМ», созданный в ИДСТУ СО РАН механизм, который создан с использованием декларативных спецификаций. Основные функции модуля администрирования:

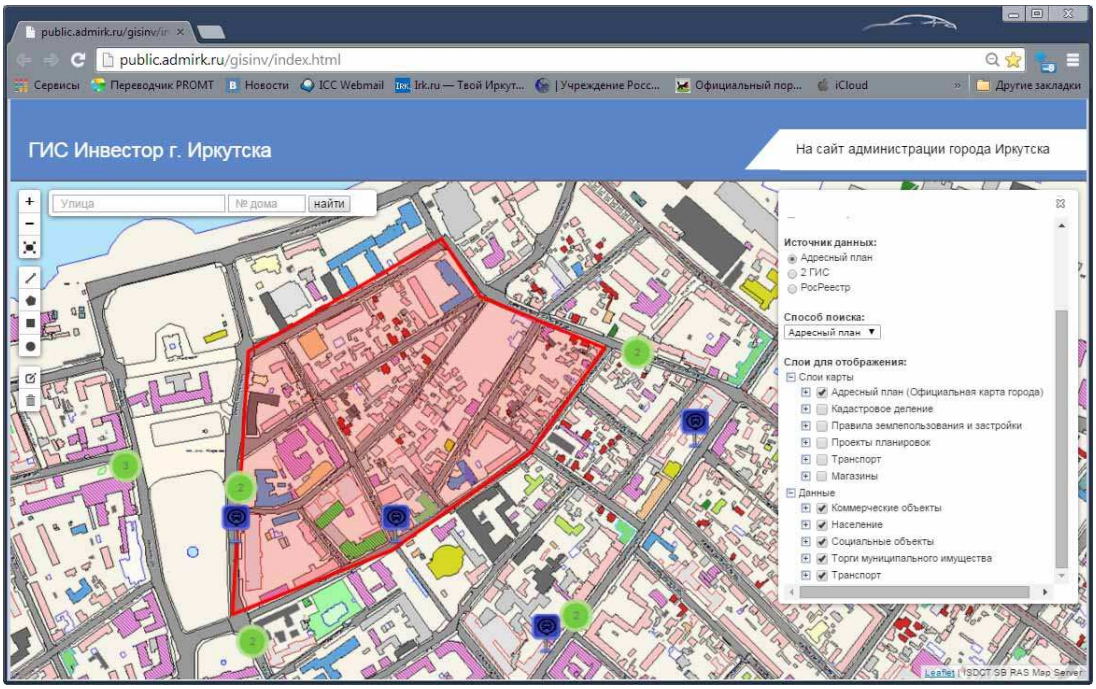

Рис. 4. Запросная система ГИС «Инвестор»

- управление логической и физической структурой графической и связанной с ней информацией тематических баз данных для публикации и управления конфигурацией созданного Интернет-ресурса;
- автоматическая загрузка данных в БД из таблиц Excel;
- редактирование табличных данных;
- загрузка новых слоев картографических данных для отображения в Web-интерфейсе;
- просмотр слоёв в подготовленном к публикации виде и фиксация определенной информации от ответственного лица, подтверждающей проведенный контроль слоя и его готовность к публикации (дата, время, реквизиты ответственного лица).

Под декларативными спецификациями БД ГИС «Инвестор» понимается метаинформация, описывающая структуру этой БД [5,6,7,8]. Декларативные спецификации используются при загрузке и отображении табличной информации с помощью подсистемы «Администрирование». При получении новой таблицы система создает метаописание таблицы в файле настроек и связывает ее с другими таблицами при необходимости. Автоматически создаются html-шаблоны на основе этих метаописаний для отображения табличных данных в Web-интерфейсе. Далее исполняемая библиотека обрабатывает эти метаописания и отображает данные в удобном для пользователя виде. Также на основе этого описания создаются определенные связи для использования информации из таблиц при создании пользовательских запросов.

В подсистеме «Администрирование» реализована функция загрузки пользовательских слоев, которая проходит в несколько этапов: пользователь выбирает новый слой в формате ГИС «КАРТА», после этого слой при помощи специального конвертора преобразуется в оригинальный формат SMD. Далее слой в формате SMD перепроецируется в проекцию WGS84. Это обеспечивает совместимость отображаемой информации со сторонними WMS сервисами (Google, OpenStreetMap и т.д.), так как их карты представлены в этой проекции.

В подсистеме «Администрирование» реализована публикация информации из слоев с ограниченным (служебным) доступом об инженерной инфраструктуре города и возможности подключения к её объектам. С помощью специального конвертора генерируется бинарный файл в формате OMR (object-relational mapping), в котором хранится информация о наличии объектов в квадрате или их отсутствии, этот файл не хранит в себе метрику и семантику. Данные из этого файла отображаются при формировании отчета по заданной области. Для этого карта разбивается на квадраты с регулярной сеткой шагом в 200 метров. По запросу пользователя формируется текущий номер квадрата и отчет в виде максимального, минимального расстояния до объектов инженерной инфраструктуры в данном квадрате. Сконвертированные слои инженерной инфраструктуры невидимы для пользователей и могут использоваться для формирования отчетов. Данная информация включается состав отчетов только, если активирован слой инженерной инфраструктуры в списке слоев ГИС «Инвестор» подсистемы Web-публикации.

ГИС «Инвестор» города Иркутска внедрен в эксплуатацию и обеспечивает простое, доступное наращивание информационной составляющей для публикации путём настройки файлов конфигурации. Это полноценный аналитический инструмент в сети Интернет, который позволяет провести комплексный анализ текстовых и картографических данных об инфраструктуре города (http://public.admirk.ru/gisinv/index.html), что значительно упрощает принятие решений об инвестиционной привлекательности объектов недвижимости.

Список литературы

1. Бычков И. В., Гаченко А. С., Ружников Г. М., Маджара Т. И., Фереферов Е. С., Хмельнов А. Е. Внедрение современных информационных технологий в региональных проектах // Вестник НГУ. – 2008. – Т. 6, № 1. – С. 15-24.
2. Фереферов Е. С., Гаченко А. С., Ружников Г. М., Хмельнов А. Е. Муниципальная информационная система обеспечения градостроительной деятельности // Вычислительные технологии. – 2008. – Т. 13, Специальный выпуск 1. – С. 11-16.
3. Гаченко А. С., Ружников Г. М., Фереферов Е. С., Хмельнов А. Е. Разработка информационной системы обеспечения градостроительной деятельности в муниципальных образованиях // Вестник НГУ. – 2008. – Т. 6, № 3. – С. 72-79.
4. Фёдоров Р. К., Бычков И. В., Хмельнов А. Е., Новицкий Ю. А., Ружников Г. М., Гаченко А. С., Фереферов Е. С., Шигаров А. О., Парамонов В. В., Попова А. К. Разработка геоинформационной системы "Адресный план"г. Иркутска // Современный технологии Системный анализ Моделирование. – 2009. – № 3 (23). – С. 14-19.
5. Бычков И. В., Ружников Г. М., Хмельнов А. Е., Фёдоров Р. К., Гаченко А. С., Шигаров А. О., Парамонов В. В. Формирование компонентов инфраструктуры пространственных данных для управления территориальным развитием // Вестник Кемеровского государственного университета. – 2012. – Т. 52, № 4. – С. 30-37.
6. Бычков И. В., Плюснин В. М., Ружников Г. М., Хмельнов А. Е., Фёдоров Р. К., Гаченко А. С. Создание инфраструктуры пространственных данных в управлении регионов // География и природные ресурсы. – 2013. – № 2. – С. 145-150.
7. Гаченко А. С., Бычков И. В., Ружников Г. М., Фёдоров Р. К., Хмельнов А. Е. Интеграционный подход формирования инфраструктуры пространственных данных // МИТ 2013 – Сербия, 2013. – С. 86.
8. Гаченко А. С., Ружников Г. М., Хмельнов А. Е. Применение ГИС и ВЕБ-технологий для создания интегрированных информационно-аналитических систем // Инфраструктура научных информационных ресурсов и систем — Москва: , 2014. – Т. I, – С. 190-200.
9. Leaflet is a modern open-source JavaScript library for mobile-friendly interactive maps [Электрон. ресурс]. –2015. – URL: http://leafletjs.com/ (дата обращения: 20.05.2015)

Организация Центра Сбора Экспериментальных Геофизических Данных в Реальном Времени для Исследования Ближнего Космоса

Ж.Ш. Жантаев[1], Б.Т. Жумабаев[2], О.Н. Крякунова[2], Н.Ф. Николаевский[2], А.М. Малимбаев[2]

[1] АО «Национальный центр космических исследований и технологий», Алматы, Казахстан
[2] ДТОО «Институт ионосферы», Алматы, Казахстан
nanozavr@mail.ru

Аннотация. В Казахстане создана многоуровневая система мониторинга и прогноза космической погоды, включающая измерения интенсивности космических лучей, напряженности геомагнитного поля и потока радиоизлучения Солнца. Вся экспериментальная информация обновляется в режиме реального времени и поступает в комплексную базу данных, архивная информация представляется в интерактивном режиме.

Ключевые слова: GPS, космические аппараты (КА), околоземное космическое пространство (ОКП).

Введение

В наше время невозможно представить жизнь человека без космических технологий. GPS-навигаторы, спутниковые телефоны и многое другое – всё это обязано своим существованием многочисленным космическим аппаратам (КА). Их назначение различно: это спутники связи, космические станции, орбитальные телескопы, спутники обеспечения систем GPS и ГЛОНАСС, метеорологические спутники и многие другие. Космическая среда, в которой работают спутники, далеко не безопасна. На каждый аппарат оказывают негативное влияние различные факторы, способные вызывать сбои и неисправности на борту спутника, а в некоторых случаях даже приводящие к потере аппарата. В целях обеспечения безопасности функционирования космических аппаратов и космических полётов необходимо учитывать это воздействие. В качестве одного из наиболее трагических примеров аварий КА по вине геофизических факторов можно привести гибель американской орбитальной станции «Скайлэб». В 1972 году мощная солнечная вспышка вызвала нагрев верхних слоёв атмосферы, увеличив тем самым её плотность. Вследствие возросшей плотности среды станция начала терять высоту слишком быстро, и на борту не хватило топлива для возврата на прежнюю орбиту. Таким образом, станция ценой в миллионы долларов была потеряна. Также радиационное излучение ведёт со временем к уменьшению КПД солнечных батарей и ухудшению оптических характеристик различных телескопов, находящихся на орбите Земли. Частицы высоких энергий вызывают сбои в бортовой электронике. В свою очередь, каждый из геофизических факторов, воздействующих на КА, зависит от солнечной активности [1]. Таким образом, функционирование КА, безусловно, связано с необходимостью контроля состояния ближнего космоса в режиме реального времени. Различные ключевые параметры, характеризующие состояние ближнего космоса, должны быть наглядно представлены в Центре представления геофизических измерений и в системе Интернет. Это даст возможность предоставлять оперативную информацию о состоянии возмущенности околоземного космического пространства, а также развивать методики прогнозирования опасных геофизических ситуаций. Задачи исследования ближнего космоса на основе современных информационных технологий признаны приоритетными и Организацией Объединенных Наций

(ООН), при которой принята Программа по решению прикладных задач в области космического пространства (http://www.unoosa.org/oosa/en/SAP/gnss/icg/iswi.html). На симпозиуме ООН (Австрия, 16-18 сентября 2013 г.) по данным, аппаратуре наблюдений и моделям космической погоды была особенно отмечена важность развития экспериментальных методов исследования ближнего космоса и создания открытых баз данных геофизических параметров околоземного космического пространства, особенно в режиме реального времени (http://www.iswi-secretariat.org). В рамках этой программы устанавливаются новые экспериментальные станции во многих точках мира для расширения возможностей исследования процессов, происходящих в ближнем космосе. К таким экспериментальных установкам относится и радиоспектрограф CALLISTO, установленный на высокогорной станции радиополигон «Орбита» ДТОО «Институт ионосферы». Измерения радиоизлучения Солнца включены в международную систему радиоспектрометров e-Callisto (http://www.e-callisto.org). Исследования космических лучей, ведущиеся с помощью высокогорного нейтронного монитора на станции космических лучей ДТОО «Институт ионосферы» входят в европейскую и мировую сеть нейтронных мониторов NMDB (www.nmdb.eu), исследования проводятся в широкой международной коллаборации специалистов в области космических лучей и космической погоды [2-4]. Таким образом, исследование физических процессов в ближнем космосе является актуальной задачей, которая проводится на современной экспериментальной базе с использованием информационных технологий и соответствует современному мировому уровню исследований [5].

Структура центра сбора экспериментальных геофизических данных в реальном времени для исследования ближнего космоса

В ДТОО «Институт ионосферы» проводится ежедневная диагностика и прогноз состояния околоземного космического пространства (ОКП) для принятия предупредительных мер против выхода из строя спутниковых и наземных технологических систем. Большое внимание уделяется усовершенствованию и развитию методик прогнозирования космической погоды. Прогноз состояния ОКП осуществляется на базе данных гелиогеофизических обсерваторий института и результатов измерений различных параметров межпланетной среды, регистрируемых на космических аппаратах. В 2009 г. в ДТОО «Институт ионосферы» создана и в настоящее время функционирует система мониторинга ОКП, которая объединяет следующие гелиогеофизические комплексы: высокогорную станцию космических лучей, радиоастрономическую и геомагнитную обсерватории (Рис. 1). Обсерватории работают в автономном круглосуточном режиме. Ежеминутно на станциях производится измерение наблюдаемых параметров ОКП, обработка данных, запись в банки архивной информации, построение графиков оперативных данных и представление результатов регистрации на сайтах станций (http://ionos.kz). Гелиогеофизическая информация в часовом режиме по FTP-каналам поступает в базу данных системы, сайт комплекса предоставляет среднечасовые оперативные и архивные результаты мониторинга.

Интегрированный банк данных представляет собой набор месячных файлов часовых значений (Рис. 2). Для решения задач исследования и прогнозирования состояния ближнего космоса, а также для проведения комплексного анализа геофизических данных необходим доступ к оперативным данным обсерваторий в реальном масштабе времени. В настоящее время проводятся работы по существенной модернизации системы и созданию центра сбора экспериментальных геофизических данных в реальном времени с минутным разрешением для исследования ближнего космоса.

Рис. 1. Комплекс гелиогеофизических обсерваторий ДТОО «Институт ионосферы»

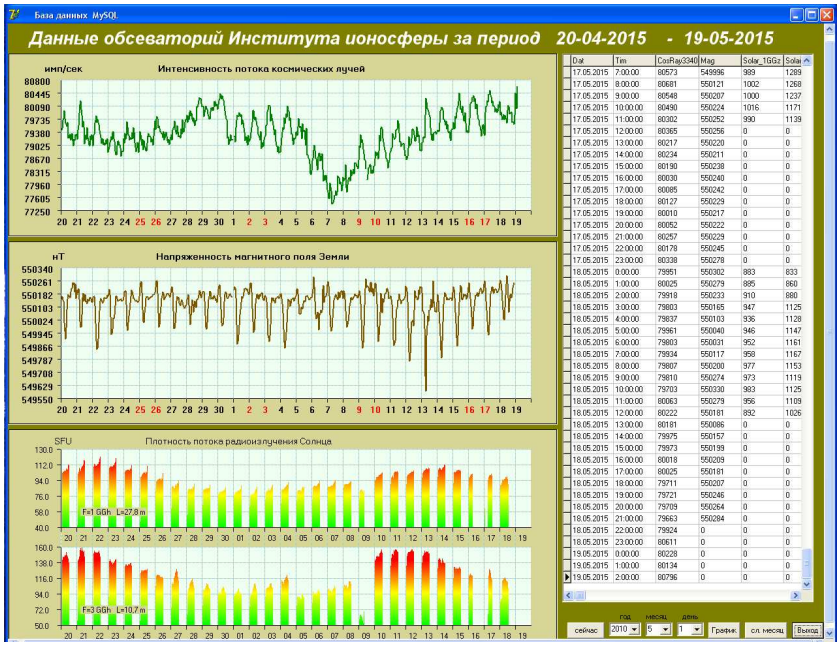

Рис. 2. Данные обсерваторий Института ионосферы (интенсивность потока космических лучей, напряженность магнитного поля Земли, плотность потока радиоизлучения Солнца) за период 20.04.2015-19.05.2015 гг.

Высокогорная станция космических лучей «ААТВ» расположена на высоте 3340 метров выше уровня моря, работает с 1973 г., на станции в автоматическом режиме на восемнадцати канальном супермониторе 18NM64 производятся ежеминутные измерения интенсивности потока космических лучей и атмосферного давления, необходимого для коррекции интенсивности космических лучей на влияние атмосферного слоя. Производится первичная обработка исходной канальной информации супермонитора 18NM64 для анализа качества, коррекции канальных данных и расчета минутных значений интенсивности космических лучей. Часовые значения рассчитываются из исходной информации с коррекцией данных каналов, параметры которых выходят за границы допустимых значений. Вся измеряемая информация записывается в базу данных и представлена на сайте станции. Передаче в центр сбора и обработки геофизической информации подлежит весь набор данных ежеминутных измерений и часовых накоплений. Геомагнитная обсерватория единственная сертифицированная обсерватория в среднеазиатском регионе, входящая в международную сеть INTERMAGNET. Данные ежедневно направляются в Международный центр данных (Ан-

глия, Эдинбург) и представляются на сайте www.gsrg.nmh.ac.uk/intermagnet, а также на сайте геомагнитной обсерватории «Алма-Ата»: http://geomag.ionos.kz. В 2010 г. обсерватория была включена в международную сеть станций по нормализации мирового Dst-индекса геомагнитной активности. Данные передаются в режиме реального времени на Dcx index server (University of Oulu, Finland) http://dcx.oulu.fi. Радиоастрономическая обсерватория расположена в горах Заилийского Алатау в 40 км от г. Алматы на высоте 2740 м выше уровня моря и функционирует с 1986 г. С 2009 г. радиоастрономическая обсерватория после технической модернизации работает в режиме регулярных измерений плотности потока радиоизлучения Солнца. Для наблюдений на двух длинах волн 30 см, 10 см используется двенадцатиметровая параболическая антенна и опорно-поворотное устройство ТНА-57М, радиометры РМ-30 и РМ-10. Получаемые данные плотности потока, измеряемые каждые пять секунд, и полученные из них часовые значения выкладываются на сайте и обновляются каждую минуту. Получаемые данные сохраняются в информационной базе. В 2011 г. на радиоастрономической обсерватории был установлен радиоспектрограф e-Callisto в комплекте с логопериодической антенной, который позволил проводить регулярные измерения радиоизлучения Солнца в диапазоне частот 45 МГц - 860 МГц. Получаемые данные с радиоспектрографа - это кадр 3600x200 точек, который каждые 15 минут сохраняется в базе данных. Данные, полученные с радиометров РМ-30, РМ-10 и на радиоспектрографе e-Callisto, подлежат передаче в центр сбора и обработки геофизической информации. Разработан общий для обсерваторий формат представления результатов измерений, что позволило создать унифицированную программу передачи данных в базу системы сбора. Для станции космических лучей выходными параметрами являются время измерения, величина атмосферного давления и интенсивность потока космических лучей; для магнитной обсерватории: время измерения, вариации компонентов напряженности магнитного поля по координатам x,y,z и величина полного вектора магнитного поля Земли; для радиоастрономической обсерватории: время измерения, плотность потока на частотах 1 ГГц и 3 ГГц. При разработке алгоритма передачи информации в систему сбора геофизической информации учитывается удаленность обсерваторий, расположение в горной труднодоступной местности, что иногда приводит к отказам связной аппаратуры и сбоям в функционировании системы Интернет. Поэтому для передачи в систему сбора всей зарегистрированной на станциях информации выходные данные представляются в виде пакета измерений и пополняют пакет при невозможности отправки данных. При разработке программы передачи предусматривается возможность последовательной выборки данных из пакета накопленных измерений. Программное обеспечение обсерваторий дополняется блоками формирования пакетов выходных данных. После первичной обработки исходной информации управляющая программа системы регистрации инициализирует блок формирования данных. Программа передачи данных извлекает информацию для передачи по каналу связи из файла и стирает его, если процесс передачи прошел успешно. Поэтому блок формирования в первую очередь проверяет, насколько успешно передана информация предыдущего сеанса. Если данные переданы, открывается новый файл и в него записываются данные последнего сеанса. Если передача данных последнего сеанса прошла не успешно, то файл с данными остается не стертым, и в него добавляется строка с данными последнего сеанса. Таким образом, обеспечивается отправка всей информации даже при сбоях и временном отсутствии связи. Программная реализация этого алгоритма предусмотрена для каждой обсерватории (Рис. 3).

Программа передачи информации в систему сбора данных реализована на языке Delphi7 с использованием библиотеки DirectMySQLObjects122 для организации связи, передачи данных по каналу и записи в базу системы сбора информации. При всех своих изменениях язык SQL остаётся единственным механизмом связи между прикладным программным

Рис. 3. Алгоритм подготовки информации к передаче в систему сбора

обеспечением и базой данных. В то же время, современные системы управления базами данных, а также информационные системы, использующие системы управления базами данных, предоставляют пользователю развитые средства визуального построения запросов. В программе используется библиотека DirectMySQLObjects122, которая предоставляет все функции, необходимые для соединения с сервером, посылки команды, выборки считанной информации. База данных дополняется новыми видами измерений: гамма-квантов вторичного космического излучения и интенсивности тепловых нейтронов (высота 3340 м над уровнем моря). На следующем этапе работ будет разработан сайт для представления оперативной и архивной информации базы центра в графическом и текстовом виде в интерактивном режиме.

Заключение

Уникальность данных геофизических обсерваторий института, обусловленная расположением на высокогорье и применением современной измерительной аппаратуры, стабильность работы в течение десятков лет, вызывает большой интерес международного научного сообщества. Казахстанские исследования, проводимые в рамках настоящей работы, включены в программу Комитета ООН по мирному исследованию космоса UNCOPUOS «Международные инициативы в области исследования космической погоды» (International Space Weather Initiative-ISWI). Создаваемая система сбора, хранения и представления данных сети обсерваторий Казахстана в режиме реального времени будет предусматривать доступ через Интернет в интерактивном режиме к оперативным и архивным данным гелиогеофизических измерений с высоким временным разрешением на сайте ДТОО «Институт ионосферы» (www.ionos.kz) как для казахстанских исследователей, так и для мирового научного сообщества.

Список литературы

1. Гальперин Ю.И., Дмитриев А.В., Зеленый Л. М., Панасюк М. И. «Влияние космической погоды на безопасность авиа и космических полетов» [Электрон. ресурс]. – URL: `http://www.kosmofizika.ru/pdf/sp_wtr.pdf`
2. Ерошенко Е.А., Белов А.В., Крякунова О.Н., Янке В.Г., База данных нейтронных мониторов в реальном времени // 1-ая Всероссийская конференция по космическим лучам: матер. междунар. науч. конф. – Москва, – МГУ. – 5-9 июля 2010 г. – СКЛ/ SCR 20.

3. Belov A.V., Drobzhev V.I., Eroshenko E.A., Kryakunova O.N., Nikolaevskiy N.F., Yanke V.G. Zhantaev Zh.Sh., Space Weather Research by means of High Mountain Alma-Ata Neutron Monitor. Multi-Wavelength Investigations of Solar Activity //Proceeding of the 23th Symposium of the International Astronomical Union. Cambridge University Press. – 2004. – P. 543-544.
4. Abunina M., Papaioannou A., Gerontidou M., Paschalis P., Abunin A., Gaidash S., Tsepakina I., Malimbayev A., Belov A., Mavromichalaki H., Kryakunova O., Velinov P. Forecasting Geomagnetic Conditions in near-Earth space //Journal of Physics: Conference Series –2013/ – P.409.– 012197. (doi:10.1088/1742-6596/409/1/012197) IOP Publishing.
5. Zhantayev Zh., Kryakunova O., Nikolayevskiy N., Zhumabayev B. Sun and Geosphere // Space weather research in Kazakhstan – Алматы, 2014. – Vol. 9 (1-2). – P.71-74.

Концепция и Архитектура Интернет–Портала для Исследования Теплофизических Свойств Металлов

Ю.И. Молородов[1], А.М. Зеленчук[2]

[1] Институт вычислительных технологий СО РАН, Новосибирск, Россия
[2] Новосибирский государственный университет, Новосибирск, Россия
yumo@ict.sbras.ru,andrey.zelenchuk@gmail.com

Аннотация. Рассматривается интернет-портал знаний, обеспечивающий интеллектуальный анализ и интеграцию знаний и информационных ресурсов по изучению теплофизических свойств материалов, в том числе органических и неорганических веществ в широком диапазоне температур. Портал обеспечивает систематизацию и интеграцию доступных информационных ресурсов в едином информационном пространстве. Использование онтологии для описания предметной области портала позволяет семантически структурировать его информационное наполнение и организовать в нем навигацию и содержательный поиск информации. Для разработки данного портала используется технология, ориентированная на экспертов предметных областей. Портал содержит систему импорта данных из файлов в базу данных (БД), которая может находиться локально или на другом сетевом ресурсе. Для отображения данных в виде интерактивных графиков, с использованием алгоритмов сглаживания, интерполяции и аппроксимации используются ресурсы разработанные на основе свободных программных продуктов.

Ключевые слова: Портал знаний, теплофизические свойства материалов, информационные ресурсы, онтология, содержательный поиск, управляемая онтологией навигация.

Введение

В последнее время в мире активно разрабатываются и используются специализированные коммерческие пакеты программ, которые позволяют получать данные по теплофизическим свойствам для некоторых классов веществ и материалов. Опыт практического использования таких программ выявил их существенные недостатки. Они не дают доступа к таблицам первичных экспериментальных данных, не позволяют выбирать модели для их обработки. Используемые модели не являются физически обоснованными, что не позволяет проводить экстраполяцию данных, отсутствует также реальная возможность прогнозирования свойств и т.д. Все это делает их пригодными только для проведения инженерных расчетов, а не научных исследований. Информационная система, которая могла-бы объединить достоинства существующих программных средств и устранить перечисленные выше недостатки, будет являться важным инструментом для развития теорий свойств веществ и фазовых превращений.

Интеграция информационных ресурсов в единую информационную среду и организация доступа к вычислительным ресурсам – это еще одно из важнейших направлений развития современных информационных технологий. Решение проблем создания и интеграции информационных ресурсов и продуктов становится необходимым условием развития многих стран, в том числе и России.

Современное состояние проблемы

Научные базы данных содержат информацию о свойствах изучаемых объектов и являются либо чисто фактографическими, либо содержат еще библиографическую и текстовую

информацию. В последние годы стали разрабатываться многосредовые (multimedia) базы данных, содержащие одновременно текст, изображение и звук. Большая часть баз данных, особенно в нашей стране, являются локализованными в лабораториях и научных центрах и малодоступны извне. За рубежом с развитием сетевых технологий в последние десять-пятнадцать лет стали создаваться национальные и международные службы научных баз данных. Например, база данных по свойствам материалов университета Пердью (штат Индиана, США) содержит фактографические базы данных по термофизическим, механическим, электрическим, оптическим и др. свойствам металлов, сплавов, композиционных материалов, полимеров. С 1988 г. она доступна в интерактивном режиме для пользователей США, работающих на персональных компьютерах. Такого рода базы данных имеются и в других странах, например, в университете Тохоку (Япония) построена фактографическая база данных по магнитным материалам, которая ориентирована на создание и использование фактографических баз данных непосредственно на рабочем месте исследователя. Там же создана фактографическая база данных по аморфным материалам, которая ведется с помощью реляционной системы управления базами данных (СУБД). Она открыта для интерактивного доступа по компьютерным сетям, а новые поступления в базу данных отображаются в системе электронных телеконференций. Однако, как правило, эти базы данных недоступны для внешних пользователей. Некоторые отечественные и многие зарубежные базы данных являются коммерческими и доступны только за довольно высокую плату. Предлагаемая в данной работе сетевая база данных по диаграммам состояния и физико-химическим свойствам материалов не имеет аналогов в России. Из зарубежных баз данных такого типа известны только коммерческие.

Обоснование необходимости проведения работы. Цели и задачи

Потребность многих исследователей в достоверной информации о диаграммах состояния материалов и их термодинамических свойствах чрезвычайно велика. Имеющийся в литературе фактический материал по этим вопросам разбросан по огромному числу периодических изданий, многие из которых недоступны по причине их отсутствия в библиотеках города, региона и даже страны. Имеющийся же справочный материал зачастую также плохо доступен либо устарел. Достаточно сказать, что в библиотеках такого крупного научного центра, как Екатеринбург, находятся справочники по диаграммам состояния, изданные, как правило, 30-40 лет назад. На момент написания статьи отсутствует даже знаменитый справочник Р.Халтгрена с соавторами (R.Hultgren, P.D.Desai, D.T.Hamkins, M.Gleiser, K.K.Kelley. Selected values of the thermodynamic properties of binary alloys. ASM. Metal Park. Ohio. 1973), не говоря уже о более поздних изданиях. Недоступны специализированные журналы, например, Bull. of Phase Diagrams, CALPHAD и т.п. Именно поэтому создание сетевых баз данных, позволяющих собрать, систематизировать научную информацию и сделать ее доступной широкому кругу исследователей, является актуальной задачей. До последнего времени создаваемые разными коллективами базы данных были доступны лишь ограниченному кругу пользователей. Это касается и поддерживаемого в Институте металлургии УрО РАН банка данных АСТРА.OWN. Однако на современном этапе именно с помощью разработки и создания сетевой базы данных на основе WWW-технологии возможно сделать собранную информацию доступной максимально широкому кругу пользователей и, таким образом, существенно повысить эффективность научно-исследовательской работы.

Кончной целью предлагаемой работы является создание фактографической сетевой мультимедийной базы данных по широкому кругу физико-химических свойств материалов в конденсированном состоянии.

Для достижения указанной целями нам нужно решить следующие задачи:

- разработать действующую модель фактографической базы данных по физико-химическим свойствам неорганических материалов;
- организовать экспертный отбор опытных данных (как полученных в результате проведения эксперимента, так и взятых из литературных источников) для наполнения базы данных;
- разместить в базе данных информацию о диаграммах состояний и физико-химических свойств материалов;
- реализовать новые возможности информационного обеспечения исследователей, решающих фундаментальные задачи в области расчетов и прогнозирования свойств многокомпонентных систем, чтобы создавать новые технологии и материалы с заданными свойствами.

Онтологический подход

Теплофизика – одна из дисциплин, в которых центральное место занимает работа с численными данными. При работе с ними приходится учитывать, что в публикациях и БД используются несколько типовых форм, а именно: табличную, графическую и математическую (в виде хранимых формул или программных кодов). Графическая форма иллюстрирует характер зависимостей, рассеяние опытных точек и т.п. Табличная форма наиболее надежна при передаче данных, легко контролируема в отношении пропусков, ошибок в знаке или порядке величины и т.п. Математическая форма, избавляя от интерполяции, требует повышенной тщательности в обнаружение ошибок, легко вылавливаемых в табличной форме. Доминирующей формой в экспериментальных работах и справочниках является именно табличная форма.

Чтобы разрабатываемый портал знаний мог предоставлять пользователям описанные выше возможности, он должен не только иметь гибкие средства представления разнородной информации и содержательного доступа к ней, но и обеспечивать оперативное управление своим информационным наполнением. Всем этим целям служит информационная модель портала знаний, описывающая предметную область.

Перспективным подходом при проектировании портала является моделирование системы с использованием онтологии [1]. Классическое определение онтологии было дано Грубером [2] и слегка модифицировано Борстом [3]: «*Онтология* – это формальная спецификация разделяемой концептуализации». Под *концептуализацией* понимается строгое описание понятий предметной области и семантических связей (отношений) между ними, в том числе средствами естественного языка. Концептуализация должна быть *разделяемой*, то есть фиксировать общие знания, которые признаны некоторой группой, а не частные знания конкретного индивидуума. Под формальной спецификацией понимается представление знаний (концептов) в формально определённом формате, например, в терминах некоторого языка, понимаемого машиной.

Таким образом, онтология позволяет описать предметную область на формальном языке, что позволяет оперировать понятиями предметной области в рамках информационной системы [4]. Именно такой подход мы используем для работы с данными о теплофизических свойствах веществ и материалов [5-6]. Разрабатываемая онтология должна достаточно полно описывать данную предметную область. Она должна служить основой для информационных систем, работающих с теплофизическими свойствами веществ и материалов.

Одной из наиболее популярных и адекватных формальных систем для описания онтологий является OWL (Web Ontology Language), основанный на логиках описаний (Description

Logics, DL). Вокруг OWL сформировано сообщество, разработаны программные интерфейсы (API), машины рассуждения и вывода (reasoners), прикладные программы.

Знания о предметной области можно разделить на два вида: первые более стабильны и постоянны, а вторые более подвержены модификациям. В соответствии с этим делением, знания, записываемые на языке OWL, подразделяются на набор терминологических аксиом (*TBox*, terminological box) и набор утверждений об индивидах (*ABox*, assertion box). Понятия «класс» и «индивид» являются базовыми в OWL. Совокупность аксиом о классах и утверждений об индивидах вместе составляют базу знаний.

Объектами исследования в теплофизике являются материалы, в том числе вещества. Материал -– объект, свойства которого могут в заметной степени определяться сферой производства: изготовитель, марка, технология, условия хранения и прочее. Вещество -– объект, свойства которого определяются его природой: стехиометрической формулой, составом, фазой и др. На рис.1 показаны основные понятия (концепты), которые мы выделяем среди данных о свойстве вещества.

Рис. 1. Базовые концепты в теплофизике.

Тогда основная запись о теплофизических данных выглядит следующим образом: «Данное вещество в данном состоянии имеет данное свойство с данным значением, полученным из данного источника».

Вещество. Существует несколько способов идентификации вещества:

1. Номер CAS (Chemical Abstracts Service).
2. Название.
3. Химическая формула.
4. Состав (для смесей).
5. Поиск с использованием классификации.

Состояние. Вещество может находиться в различных состояниях:

1. Однофазные состояния.
   Фазы (поля равновесия):
   - Твёрдые (отличаются ориентационным и трансляционным порядком).
   - Жидкости (отличаются ориентационным порядком).
   - Газ.
   - Плазма.

2. Двухфазные состояния.
   Межфазные границы (линии равновесия)
   - «жидкость-газ» – кривая насыщения (кипения).
   - «твёрдое-газ» – кривая сублимации.
   - «твердое-жидкость» – кривая плавления.
   - «твёрдое-твёрдое».
   - «жидкость-жидкость».
3. Трёхфазные состояния.
   Особые точки (точки равновесия)
   - тройные точки;
   - критическая точка.

В разных состояниях одно и то же вещество может иметь различные наборы свойств и различные значения свойств. Например, в состоянии на кривой плавления вещество обладает свойствами твёрдого тела, свойствами жидкости и свойствами, присущими ему только на кривой плавления.

Данные. Значения свойств веществ:

1. Константа.
2. Функция одной переменной
   . аналитическая;
   . табличная (одномерная).
3. Функция двух переменных
   . аналитическая;
   . табличная (двумерная).

Источник. Источником данных в теплофизике могут служить:

1. Базы данных.
2. Статьи.
3. Собственные эксперименты.

Чаще всего свойство зависит от температуры и измеряется в ходе эксперимента при различной температуре. В результате получается одномерная таблица. Если свойство зависит ещё и от давления, и измеряется при различных давлениях, то получается двумерная таблица.

Архитектура информационной системы

При реализации информационной системы (ИС) используется многоуровневая архитектура клиент–сервер. По принципу разделения ответственностей бизнес–логика выносится в отдельный модуль. Графический веб–интерфейс (Web UI) взаимодействует с основным приложением (сервером) через программный интерфейс (API) по протоколу HTTP. В качестве формата запросов и ответом выбран формат JSON. Другие приложения также могут взаимодействовать с сервером через данный программный интерфейс.

Все числовые данные, описывающие свойства материалов хранятся в нереляционной базе данных MongoDB. При математической обработке числовых данных могут понадобиться библиотеки алгоритмов анализа. Для этого нам необходима интеграция с пакетом прикладных математических программ, в частности с пакетом Scilab [7].

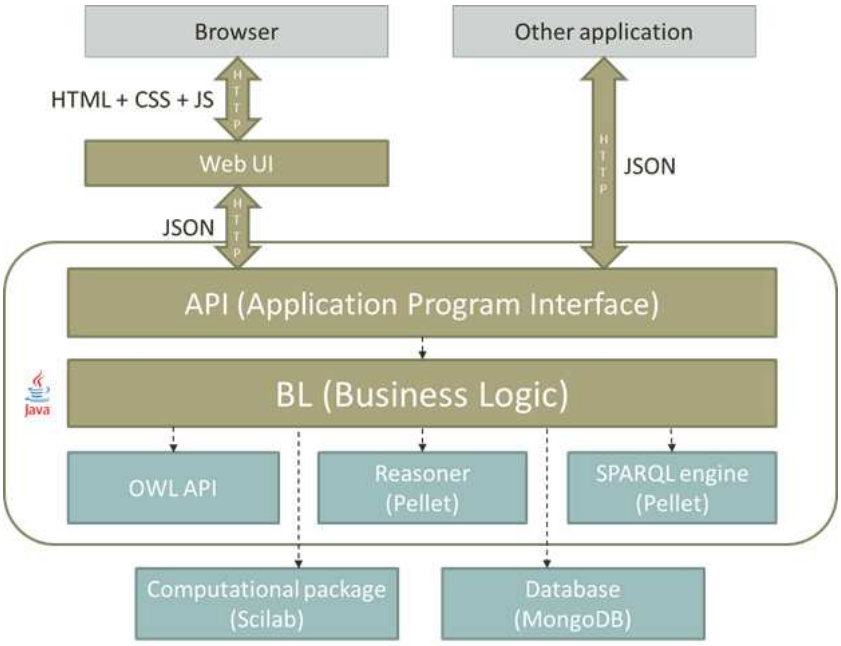

Рис. 2. Архитектура ИС.

Функциональные возможности ИС

Важной компонентой информационной системы является графический интерфейс пользователя. Так с его помощью отображается классификация материалов (рис. 3). Число около узла дерева классификатора показывает количество материалов–индивидов, относящихся непосредственно к данному классу.

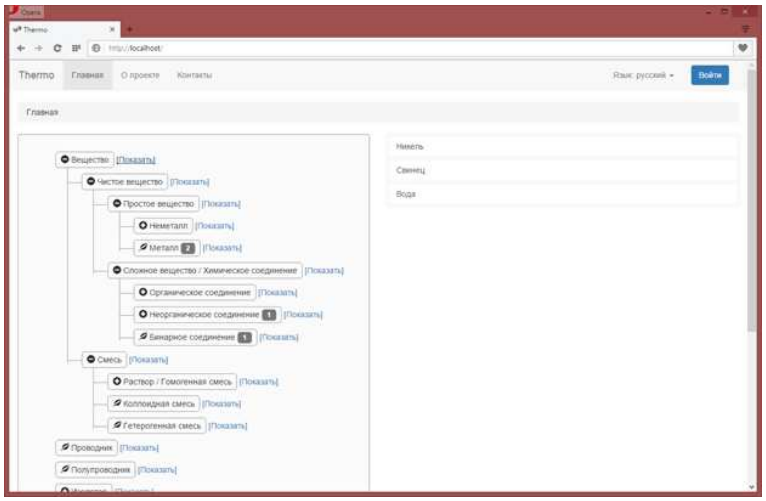

Рис. 3. Классификация материалов.

На рис.4 представлена фазовая диаграмма воды для возможных ее состояний. Она строится автоматически, исходя из имеющихся данных.

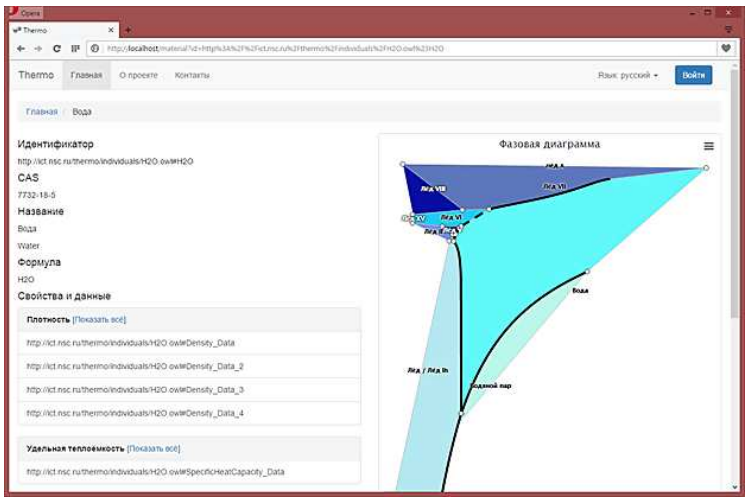

Рис. 4. Фазовая диаграмма воды H_2O.

На рис.5 представлен результат обработки данных эксперимента по исследованию зависимости плотности Никеля от температуры. Рассмотрены все фазы состояний от твердой фаз до плавления, включая фазовый переход (ФП) мы не можеи обработать всю совоокупность данных (из-за наличия ФП), поэтому мы обрабатываем участки с гладкими кривыми. Каждый из них может обрабатываться соответствующим алгоритмом аппроксимации: линейной, квадратичной или кубической. В качестве функций могут использоваться полиномы, сплайн-функции или кривые Безье.

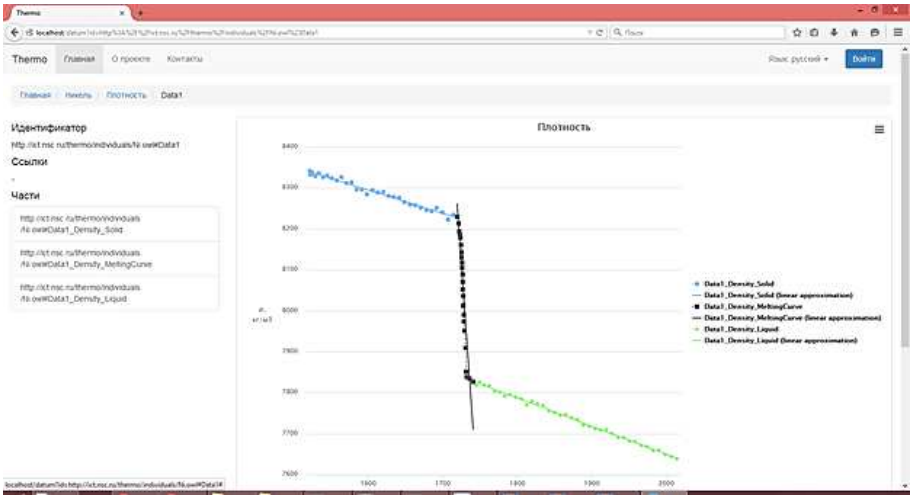

Рис. 5. Динамика изменения плотности Ni от температуры.

На одном графике можно отобразить:

- все имеющиеся данные о свойстве (из всех источников);
- отдельные данные из одного источника;
- часть данных, относящуюся к одному состоянию вещества.

Заключение

Рассмотренный портал знаний обеспечивает систематизацию и интеграцию знаний и доступных информационных ресурсов, относящихся к изучению свойств органических и неорганических соединений в единое информационное пространство, и содержательный доступ к ним. Систематизация и структуризация таких знаний и информационных ресурсов выполнена на основе онтологии, поэтому доступ к ним осуществляется путем навигации по дереву понятий онтологии и контенту портала и через средства поиска в терминах его предметной области.

Реализована возможность импорта данные из файлов, их просматр и редактирование. Результат обработки может быть представлен как в виде таблиц так и в виде графиков, построенных на основе алгоритмов сглаживания, интерполяции и аппроксимации.

Список литературы

1. Erkimbaev A.O., Zhizhchenko A.B., Zitserman V.Y., Kobzev G.A., Son E.E., and Sotnikov A.N. Integration of Databases on Substance Properties: // Automatic Documentation and Mathematical Linguistics. 2012. Vol. 46. No. 4. pp. 170–176.
2. Gruber T.R. Toward principles for the design of ontologies used for knowledge sharing // Human-Computer Studies, No. 43, 1993. pp. 907-928.
3. Borst W.N. Construction of engineering ontologies for knowledge sharing and reuse // CTIT Ph. D-thesis series No. 97-14. Universiteit Twente. Enschede. 1997. 243 pp.
4. Загорулько Ю.А., Боровикова О.И. Информационная модель портала научных знаний // Информационные технологии. – 2009. – № 12. –С. 2–7.
5. Загорулько Г. Б., Молородов Ю. И., Федотов А. М. Систематизация знаний по теплофизическим свойствам веществ // Вестник НГУ. Серия: Информационные технологии. - 2014. - Том 12, Выпуск № 3. - С. 48-56. - ISSN 1818-7900.
6. В.Б. Барахнин, Ю.И. Молородов, С.В. Станкус, А.М. Федотов Информационные технологии для задач теплофизических свойств веществ. //Информатика и системы управления. Автоматизированные системы и комплексы. 2013, № 4(38).- С. 149-157
7. Open source software for numerical computation. [Электрон. ресурс]. – URL: `http://www.scilab.org/`

Оптимизация Непрерывных Линейных Систем с Ограниченным Управлением

З.Н. Мурзабеков

НИИ Математики и механики Казахский Национальный Университет им. аль-Фараби, Алматы, Казахстан
murzabekov-zein@mail.ru

Аннотация. Рассматривается задача оптимального управления нестационарными линейными системами со свободными правыми концами траекторий. Квадратичный целевой функционал зависит от управления и состояния объекта. Предлагается новый подход конструирования регулятора для систем автоматического управления с учетом ограничений на значения управления. Задача решена с использованием множителей Лагранжа специального вида.

Ключевые слова: непрерывная линейная система, ограниченное управление, квадратичный целевой функционал, задача оптимального управления, метод множителей Лагранжа.

1 Введение

В работах в области автоматического управления можно найти различные примеры математической постановки и методы решения задач оптимального управления. Впервые аналитическое решение линейно-квадратичной задачи без ограничений на управление и со свободными правыми концами траекторий было получено в работах А.М. Летова и Р.Е. Калмана [1], [2]. В основе решения задач оптимального управления при построении программного управления лежит принцип максимума Л.С. Понтрягина [3]. Решение этой же задачи в виде синтеза оптимального управления с обратной связью основано на методе динамического программирования Р. Беллмана [4] или использованы достаточные условия оптимальности В.Ф. Кротова [5] и принцип расширения В.И. Гурмана [6].

Различные математические постановки задач оптимального управления, а также их классификация приведены в [7], [8]. Обстоятельный обзор моделей и методов, используемых в современной теории оптимального управления, содержатся в работах [9], [10]. В работе [10] обсуждаются многие задачи, которые ставились и исследовались на протяжении всей истории теории экстремума. Главное внимание уделено принципу Лагранжа для необходимых и достаточных условий. Рассмотрены аналитические и численные методы решения задач поиска экстремума для функций многих переменных на основе необходимых и достаточных условий в [11]. Задачам оптимального управления посвящено большое количество работ теоретического и прикладного характера [12]-[22]. В работе [21] исследуется линейно-квадратичная задача оптимального управления динамическими системами с выпуклым минимизируемым функционалом и ограничениями на управление. Эта задача представляет постоянный интерес для теории оптимального управления и имеет множество приложений. В случае отсутствия ограничений получены алгоритмы синтеза оптимального управления в виде программного управления и в виде управления с обратной связью. В работе [21] отмечено, что в общем случае регулярные алгоритмы решения линейно-квадратичных задач при наличии различных ограничений отсутствуют.

Разработка различных способов построения регуляторов и соответствующих алгоритмов управления, обладающих необходимыми для приложений свойствами с учетом ограниченных управляющих воздействий, является актуальной задачей современных информационных технологий.

В данной работе рассматривается задача оптимального управления со свободными правыми концами траекторий при наличии ограничений на значения управления. Предлагается новый подход конструирования регулятора для систем автоматического управления и соответствующий алгоритм управления, основанный на принципе обратной связи, с учетом ограничений на значения управлений.

2 Постановка задачи

Пусть управляемая линейная нестационарная система описывается векторным дифференциальным уравнением:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + f(t), \quad t \in [t_0, T], \quad x(t_0) = x_0, \tag{1}$$

$$\begin{gathered} u(t) \in U(t) = \{u \mid \alpha(t) \le u(t) \le \beta(t), \ \ t \in [t_0, T]; \\ \alpha(\cdot),\ \beta(\cdot) \in C[t_0, T]\} \subset L_2([t_0, T]), R^m), \end{gathered} \tag{2}$$

где $x(t)$ – n-вектор состояния объекта управления; $u(t)$ – m-вектор кусочно-непрерывных управляющих воздействий; $A(t)$, $B(t)$ – матрицы размерностей $(n \times n)$, $(n \times m)$ соответственно; $f(t)$, $\alpha(t)$, $\beta(t)$ – непрерывные ограниченные функции; x_0 – заданный вектор; t_0, T – заданные начальный и конечный моменты времени.

Будем предполагать, что система (1) управляема. Через $\Delta(t_0, T)$ обозначим множество всех допустимых управлений, удовлетворяющих условию (2), и соответствующих траекторий $x(t, u)$ системы (1), определенных на отрезке $t_0 \le t \le T$, т.е. множество всех допустимых пар $(x(t), u(t))$:

$$\begin{gathered} \Delta(t_0, T) = \{(x, u) \mid u(t) \in U(t), \ \ \dot{x}(t) = A(t)x(t) + B(t)u(t) + f(t), \\ t \in [t_0, T], \ \ x(t_0) = x_0\}. \end{gathered}$$

Пусть на множестве $\Delta(t_0, T)$ задан функционал, который зависит от состояния объекта x и управления u:

$$J(x, u) = \frac{1}{2}x^*(T)Fx(T) + \frac{1}{2}\int_{t_0}^{T} [x^*(t)Q(t)x(t) + u^*(t)R(t)u(t)]\, dt. \tag{3}$$

Здесь $R(t)$, F — симметрические положительно определенные матрицы; $Q(t)$ — симметрическая положительно полуопределенная матрица. Символ $(^*)$ означает операцию транспонирования.

Требуется найти синтезирующее управление $\tilde{u}(x, t)$ такое, что соответствующая ему пара $(\tilde{x}(t), \tilde{u}(t))$ доставляет минимальное значение функционалу (3), где $\tilde{x}(t)$ является решением дифференциального уравнения (1) при управлении $\tilde{u}(t) = u(\tilde{x}(t), t)$, удовлетворяющем ограничению (2).

В теории автоматического управления широко распространены различные способы формирования управления. Можно реализовать позиционное управление, основанное на принципе обратной связи. Могут использоваться пропорционально-интегральные или пропорционально–интегрально-дифференциальные управления. Во многих случаях, достаточно ввести расширенный вектор состояния системы и произвести соответствующие преобразования управляющих воздействий, и тогда возникает стандартная задача оптимального управления вида (1)-(3). Отметим, что задача проектирования состоит в выборе способа

управления, при котором переходные процессы в системе удовлетворяют характеристикам качества и ограничениям.

Для решения задачи оптимального управления (1)-(3) использован метод, основанный на применении множителей Лагранжа специального вида [23], [24].

3 Решение задачи

Для решения поставленной задачи прибавим к выражению для функционала (3) систему дифференциальных уравнений (1) с множителем $\lambda_0(x,t) = K(t)x+q(t)$, а также следующее выражение

$$\lambda_1^*(t)[\alpha(t) - u(t)] + \lambda_2^*(t)[u(t) - \beta(t)] + \lambda_3^*(t)[x(t) - W^{-1}(t)q(t) - z(t)],$$

где $\lambda_1(t) \geq 0$, $\lambda_2(t) \geq 0$. В результате получим следующий функционал:

$$\begin{gathered} L(x,u) = \frac{1}{2}x^*(T)Fx(T) + \int_{t_0}^{T}\Big\{\frac{1}{2}x^*(t)Q(t)x(t) + \frac{1}{2}u^*(t)R(t)u(t)+ \\ +[K(t)x(t)+q(t)]^*[A(t)x(t)+B(t)u(t)+f(t)-\dot{x}(t)]+ \\ +\lambda_1^*(t)[\alpha(t)-u(t)] + \lambda_2^*(t)[u(t)-\beta(t)]+ \\ +\lambda_3^*(t)[x(t)-W^{-1}(t)q(t)-z(t)]\Big\}dt, \end{gathered} \tag{4}$$

где $q(t)$ – n-вектор; $K(t)$ – $(n \times n)$-матрица. Множитель $\lambda_0(x,t) = K(t)x + q(t)$ снимает ограничения, налагаемые на $(x(t), u(t))$, в виде системы дифференциальных уравнений (1), а функции $\lambda_1(t)$, $\lambda_2(t)$ – соответствующие ограничения, налагаемые на управление (2); множитель $\lambda_3(t)$ сохраняет свойства системы для определения граничных условий. Такое представление функционала (4) позволяет свести исходную задачу на условный экстремум к задаче на безусловный экстремум.

Для преобразования функционала (4) введем в рассмотрение функции:

$$V(x,t) = \frac{1}{2}x^*K(t)x + x^*q(t), \tag{5}$$

$$\begin{gathered} M(x,u,t) = \frac{1}{2}x^*Q(t)x + \frac{1}{2}u^*R(t)u + [K(t)x+q(t)]^*[A(t)x+ \\ +B(t)u+f(t)] + \frac{1}{2}x^*\dot{K}(t)x + x^*\dot{q}(t) + \lambda_1^*(t)[\alpha(t)-u(t)]+ \\ +\lambda_2^*(t)[u(t)-\beta(t)] + \lambda_3^*(t)[x - W^{-1}(t)q(t) - z(t)], \end{gathered} \tag{6}$$

$$V_1(x(T),T) = \frac{1}{2}x^*(T)Fx(T) - V(x(T),T), \tag{7}$$

Теперь функционал (4), с учетом введенных функций (5)-(7), можно представить в следующем виде:

$$L(x,u) = V(x(t_0),t_0) + V(x(T),T) + \int_{t_0}^{T} M(x(t),u(t),t)\,dt. \tag{8}$$

Выберем управление u таким образом, чтобы функция $M(x,u,t)$ достигала минимума по u и x для любого $t \in [t_0,T]$. Необходимыми условиями минимума функции $M(x,u,t)$ являются следующие соотношения:

$$\frac{\partial M(x,u,t)}{\partial u} = R(t)u + B^*(t)[K(t)x+q(t)] - \lambda_1(t) + \lambda_2(t) = 0, \tag{9}$$

$$\frac{\partial M(x,u,t)}{\partial x} = Q(t)x + K(t)[A(t)x + B(t)u + f(t)] + \\ +A^*(t)[K(t)x + q(t)] + \dot{K}(t)x + \dot{q}(t) + \lambda_3(t) = 0, \tag{10}$$

$$\frac{\partial V_1(x(T),T)}{\partial x(T)} = [F - K(T)]x(T) - q(T) = 0. \tag{11}$$

Из (9) находим:

$$u = -R^{-1}(t)B^*(t)[K(t)x + q(t)] + R^{-1}(t)[\lambda_1(t) - \lambda_2(t)]. \tag{12}$$

Условие (10) с учетом управления (12) можно записать в виде

$$\frac{\partial M(x,u,t)}{\partial x} = [\dot{K}(t) + A^*(t)K(t) + K(t)A(t) - K(t)B_1(t)K(t) + \\ +Q(t)]x + \dot{q}(t) + A_1^*(t)q(t) + K(t)f(t) + \\ +K(t)B(t)R^{-1}(t)[\lambda_1(t) - \lambda_2(t)] + \lambda_3(t) = 0,$$

где $A_1(t) = A(t) - B_1(t)K(t)$, $B_1(t) = B(t)R^{-1}(t)B^*(t)$.

Пусть матрица $K(t)$ является решением уравнения Риккати

$$\dot{K}(t) = -A^*(t)K(t) - K(t)A(t) + K(t)B_1(t)K(t) - Q(t), \\ K(T) = K_T, \tag{13}$$

а вектор $q(t)$ удовлетворяет дифференциальному уравнению

$$\dot{q}(t) = -A_1^*(t)q(t) - K(t)f(t) + W(t)B(t)\varphi(t), \tag{14}$$

где $\varphi(t) = R^{-1}(t)[\lambda_1(t) - \lambda_2(t)]$, а множитель $\lambda_3(t)$ равен

$$\lambda_3(t) = -[K(t) + W(t)]B(t)\varphi(t). \tag{15}$$

Тогда выполняются условия стационарности (9)-(11), а дифференциальное уравнение, определяющее закон движения системы (1) с управлением (12), будет следующим:

$$\dot{x}(t) = A_1(t)x(t) - B_1(t)q(t) + B(t)\varphi(t) + f(t), \quad x(t_0) = x_0. \tag{16}$$

Отметим, что определение начального условия для дифференциального уравнения (14) целесообразно осуществить из следующего соотношения:

$$q(t) = W(t)[x(t) - z(t)], \quad t \in [t_0, T], \tag{17}$$

где матрица $W(t)$ удовлетворяет дифференциальному уравнению

$$\dot{W}(t) = -A_1^*(t)W(t) - W(t)A_1(t) + W(t)B_1(t)W(t), \quad W(T) = [F - K_T], \tag{18}$$

а вектор $z(t)$ является решением дифференциального уравнения

$$\dot{z}(t) = A_1(t)z(t) + [E + W^{-1}(t)K(t)]f(t), \quad z(T) = 0, \tag{19}$$

где E – единичная матрица.

Отсюда, используя решения уравнений (18) и (19), из соотношения (17) можно определить начальное условие для уравнения (14):

$$q(t_0) = W(t_0)[x(t_0) - z(t_0)]. \tag{20}$$

Теперь обозначим через

$$w(x,t) = -R^{-1}(t)B^*(t)[K(t)x + q(t)] \tag{21}$$

и определим множители $\lambda_1(t) \geq 0$, $\lambda_2(t) \geq 0$, управление $\tilde{u}(t)$ так, чтобы выполнялись условия:

$$R(t)[\tilde{u}(t) - w(x,t)] - \lambda_1(t) + \lambda_2(t) = 0, \tag{22}$$

$$\lambda_1^*(t)[\alpha(t) - \tilde{u}(t)] = 0, \quad \lambda_2^*(t)[\tilde{u}(t) - \beta(t)] = 0. \tag{23}$$

Для этого вычислим значения следующих векторов:

$$\begin{aligned} \lambda_1^0(t) &= R(t)\max\{0,\ \alpha(t) - w(x(t),t)\}, \\ \lambda_2^0(t) &= R(t)\max\{0,\ w(x(t),t) - \beta(t)\}. \end{aligned} \tag{24}$$

Используя (24), выберем компоненты $\lambda_{1i(t)}$, $\lambda_{2i}(t)$, $\tilde{u}_i(t)$, $(i = \overline{1,m})$ искомых векторов следующим образом:

– если $\lambda_{1i}^0(t) = 0$, $\lambda_{2i}^0(t) = 0$, то принимаем $\lambda_{1i}(t) = 0$, $\lambda_{2i}(t) = 0$, а значения управления $\tilde{u}_i(t)$ определяем из уравнения (22);

– если $\lambda_{1i}^0(t) > 0$, $\lambda_{2i}^0(t) = 0$, то принимаем $\lambda_{2i}(t) = 0$, значения $\lambda_{1i}(t)$ определяем из уравнения (22), а значения управления принимаем равными $\tilde{u}_i(t) = \alpha_i(t)$;

– если $\lambda_{1i}^0(t) = 0$, $\lambda_{2i}^0(t) > 0$, то принимаем $\lambda_{1i}(t) = 0$, значения $\lambda_{2i}(t)$ определяем из уравнения (22), а значения управления принимаем равными $\tilde{u}_i(t) = \beta_i(t)$.

Полученные результаты можно сформулировать в виде следующей теоремы.

Теорема. *Для оптимальности пары $(\tilde{x}(t), \tilde{u}(t)) \in \Delta(t_0, T)$ необходимо и достаточно, чтобы траектория $\tilde{x}(t)$ удовлетворяла дифференциальному уравнению:*

$$\dot{x}(t) = A_1(t)x(t) - B_1(t)q(t) + B(t)\varphi(t) + f(t), \quad x(t_0) = x_0,$$

а управление $\tilde{u}(t)$ определялось следующим образом:

$$\tilde{u}(t) = -R^{-1}(t)B^*(t)[K(t)\tilde{x}(t) + q(t)] + \varphi(t),$$

где матрица $K(t)$ является решением уравнения (13), вектор $q(t)$ удовлетворяет уравнению (14), а функция $\varphi(t)$ определяется следующим образом:

$$\varphi(t) = R^{-1}(t)[\lambda_1(t) - \lambda_2(t)].$$

Доказательство теоремы производится аналогично [24].

Алгоритм решения задачи. Опишем удобный для реализации на компьютере алгоритм решения задачи оптимального управления (1)-(3).

1. Проинтегрировать систему дифференциальных уравнений (13), (18), (19) для определения матриц $K(t)$, $W(t)$, вектора $z(t)$ в интервале $[t_0, T]$ при конечных условиях $K(T) = K_T$, $W(T) = [F - K_T]$, $z(T) = 0$; найти $K(t_0)$, $W(t_0)$, $z(t_0)$.

2. Задать начальное условие $x(t_0) = x_0$; вычислить $q(t_0)$ по формуле (20).

3. Проинтегрировать систему дифференциальных уравнений (13), (18), (16), (14) в интервале $[t_0, T]$ при начальных условиях $K(t_0)$, $W(t_0)$, $x(t_0)$, $q(t_0)$. В процессе интегрирования определяются оптимальная траектория $\tilde{x}(t)$ и оптимальное управление $\tilde{u}(t)$.

Если необходимо произвести расчеты для нового значения начального состояния объекта $x(t_0) = x_0$, то повторить пункты 2 и 3.

4 Пример

В качестве примера исследуем задачу оптимального управления смесителем двух потоков жидкостей, расходы которых регулируются с помощью клапанов [25]. Расходы входных потоков обозначим через $F_1(t)$ и $F_2(t)$. Жидкости имеют постоянные концентрации c_1 и c_2 растворенного в них вещества. Перемешивание жидкостей происходит в резервуаре на дне которого имеется выходное отверстие для слива смеси. Величина выходного потока $F(t)$ зависит от $h(t)$ – высоты смеси в резервуаре. Еще одной выходной переменной является концентрация полученной смеси $c(t)$.

Пусть в установившемся процессе входные и выходные переменные имеют следующие значения: $F_1(t) = F_{10}$, $F_2(t) = F_{20}$, $h(t) = h_0$, $c(t) = c_0$. Отклонения от установившихся значений обозначим через $u_1(t) = F_1(t) - F_{10}$, $u_2(t) = F_2(t) - F_{20}$, $x_1(t) = h(t) - h_0$, $x_2(t) = c(t) - c_0$. Тогда линеаризованные уравнения процесса могут быть записаны в виде

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \tag{25}$$

где $x(t) = (x_1(t), x_2(t))^*$ – вектор состояния объекта, $u(t) = (u_1(t), u_2(t))^*$ – вектор управления. В рассматриваемом примере матрицы $A(t)$ и $B(t)$ являются стационарными и имеют вид:

$$\begin{gathered} A(t) = A = \begin{pmatrix} -\dfrac{F_{10}+F_{20}}{2h_0S} & 0 \\ 0 & -\dfrac{F_{10}+F_{20}}{h_0S} \end{pmatrix}, \\ B(t) = B = \begin{pmatrix} \dfrac{1}{S} & \dfrac{1}{S} \\ \dfrac{c_1 - c_0}{h_0S} & \dfrac{c_2 - c_0}{h_0S} \end{pmatrix}, \quad t \in [t_0, T], \end{gathered} \tag{26}$$

где S – площадь поперечного сечения резервуара. На выходе системы контролируются расход смеси $F(t)$ и ее концентрация $c(t)$; желаемые значения этих переменных задаются с помощью параметров F_0 и c_0. Расход выходного потока зависит от скорости истечения смеси через отверстие на дне резервуара, которая в свою очередь зависит от высоты уровня смеси в резервуаре [26]. Высота смеси в резервуаре h_0, соответствующая желаемому расходу выходного потока F_0, может быть вычислена по формуле: $h_0 = F_0^2/(2s^2g)$, где s – площадь поперечного сечения отверстия на дне резервуара, g – ускорение свободного падения. В установившемся режиме выполнено условие баланса входных и выходного потоков: $F_{10} + F_{20} = F_0$.

Пусть требуется минимизировать квадратичный целевой функционал вида (3), где

$$F = \begin{pmatrix} 5000 & 0 \\ 0 & 10000 \end{pmatrix}, \quad Q(t) = Q = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.04 \end{pmatrix}, \quad R(t) = R = \begin{pmatrix} 20 & 0 \\ 0 & 2 \end{pmatrix}.$$

Были проведены численные расчеты при следующих значениях параметров: $F_{10} = 0.015$ м^3/с, $F_{20} = 0.005$ м^3/с, $F_0 = 0.02$ м^3/с, $c_1 = 1$ кмоль/м^3, $c_2 = 2$ кмоль/м^3, $c_0 = 1.25$ кмоль/м^3, $h_0 = 1$ м, $S = 1$ м^2, $s = 0.45 \cdot 10^{-4}$ м^2. При этих значениях параметров матрицы (26) будут иметь вид

$$A = \begin{pmatrix} -0.01 & 0 \\ 0 & -0.02 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ -0.25 & 0.75 \end{pmatrix}.$$

Начальные значения фазовых переменных были взяты равными $x_1(0) = -0.01$ м, $x_2(0) = 0.2$ кмоль/м^3. Ограничения на значения управлений заданы в виде

$$|u_1(t)| \le 0.005 \text{ м}^3/\text{с}, \quad |u_2(t)| \le 0.005 \text{ м}^3/\text{с} \tag{27}$$

на отрезке времени $t \in [0, T]$, где $T = 50$ с.

Таким образом, имеем задачу оптимального управления линейной системой второго порядка (25) с двумерным управлением с ограничениями вида (27). Квадратичный целевой функционал (3) представлен в виде функционала Больца с терминальным и интегральным слагаемыми. Правые концы траекторий системы являются свободными, «штраф» за отклонение фазовых переменных от желаемых значений в конечный момент времени задается в виде терминального слагаемого в целевом функционале.

Поскольку в рассматриваемом примере задача оптимального управления является стационарной, здесь для дифференциального уравнения Риккати (13) имеется возможность выбора такого конечного условия $K(T) = K_T$, чтобы матрица усиления $K(t)$ в синтезирующем управлении (12) была постоянной, т.е. $K(t) \equiv K_T$, $t \in [t_0, T]$. Решая алгебраическое уравнение Риккати

$$A^* K_T - K_T A + K_T B R^{-1} B^* K_T - Q = 0,$$

находим

$$K_T = \begin{pmatrix} 0.58566501 & -0.29443132 \\ -0.29443132 & 0.56192140 \end{pmatrix}.$$

Результаты численных расчетов представлены на рис. 1 и 2.

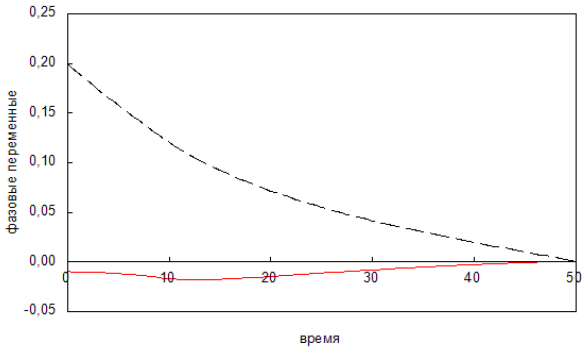

Рис. 1. Графики оптимальных траекторий
(—— $x_1(t)$, – – – $x_2(t)$)

Как видно из графиков, найденные оптимальные управления $u_1(t)$ и $u_2(t)$ удовлетворяют ограничениям (27) и обеспечивают стабилизацию системы на конечном интервале времени. В численных расчетах получены следующие конечные значения фазовых переменных: $x_1(T) \approx -0.59312 \cdot 10^{-5}$ м, $x_2(T) \approx 0.4538 \cdot 10^{-5}$ кмоль/м^3.

5 Заключение

В данной работе предложен конструктивный алгоритм построения регулятора, основанного на принципе обратной связи при наличии ограничений на значения управления. Задача

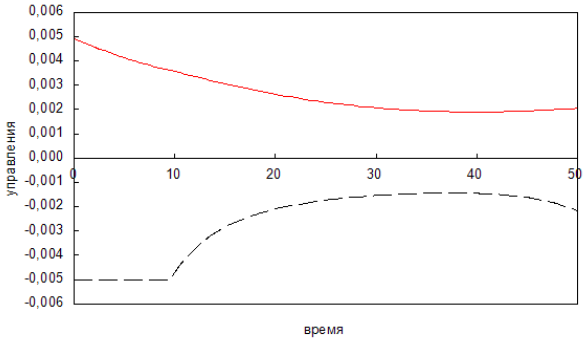

Рис. 2. Графики оптимальных управлений
(—— $u_1(t)$, – – – $u_2(t)$)

решена с использованием множителей Лагранжа специального вида, зависящих от фазовых координат и времени. За счет выбора функции $\lambda_0(x,t) = K(t)x + q(t)$ удается построить оптимальное управление по принципу обратной связи, а множители $\lambda_1(t) \geq 0$ и $\lambda_2(t) \geq 0$ выбираются таким образом, чтобы были выполнены условия дополняющей нежесткости в методе множителей Лагранжа.

Предлагаемый подход имеет преимущества при построении управления с обратной связью для стационарных систем: можно задать конечное условие $K(T) = K_T$ таким образом, чтобы свойство стационарности сохранялось и для оптимальной системы управления. Рассмотрен пример оптимальной стабилизации системы на конечном интервале времени при наличии ограничений на управление.

Список литературы

1. Летов А.М. Аналитическое конструирование регуляторов. I // Автоматика и телемеханика. 1960. – Т. 21, № 4. – С. 436-441.
2. Калман Р.Е. Об общей теории систем управления // Труды I Конгресса Международной федерации по автоматическому управлению. – Т. 2. – М.: АН СССР, 1961. – С. 521-547.
3. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. – М.: Наука, 1983. – 392 с.
4. Беллман Р., Калаба Р. Динамическое программирование и современная теория управления. – М.: Наука, 1968. – 446 с.
5. Кротов В.Ф., Гурман В.И. Методы и задачи оптимального управления. – М.: Наука, 1973. – 446 с.
6. Гурман В.И. Принцип расширения в задачах управления. – М.: Наука, 1997. – 288 с.
7. Брайсон А., Хо Ю-ши. Прикладная теория оптимального управления. – М.: Мир, 1972. – 544 с.
8. Атанс М., Фалб П. Оптимальное управление: введение в теорию и приложения. – М.: Машиностроение, 1968. – 764 с.
9. Методы классической и современной теории автоматического управления / Под ред. К.А. Пупкова, Н.Д. Егупова. – Т. 5. Методы современной теории автоматического управления. – М.: МГТУ им. Н.Э. Баумана, 2004. – 784 с.
10. Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. – М.: Физмаилит, 2005. – 384 с.
11. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах. – М., Высшая школа, 2005. – 544 с.
12. Васильев С.Н., Жерлов А.К., Федосов Е.А., Федунов Б.Е. Интеллектное управление динамическими системами. – М.: Физматлит, 2000. – 352 с.
13. Летов А.М. Динамика полета и управление. – М.: Наука, 1969. – 359 с.
14. Красовский Н.Н. Теория управления движением. – М.: Наука, 1968. – 476 с.
15. Васильев Ф.П. Численные меты решения экстремальных задач. – М.: Наука, 2002. – 520 с.
16. Лионс Ж.-Л. Оптимальное управление системами, описываемыми уравнениями с частными производными. – М.: Мир, 1972. – 414 с.

17. Бутковский А.Г. Методы управления системами с распределенными параметрами. – М.: Наука, 1975. – 568 с.
18. Егоров А.И. Оптимальное управление тепловыми и диффузионными процессами. – М.: Наука, 1978. – 464 с.
19. Ройтенберг Я.Н. Автоматическое управление. – М.: Наука, 1978. – 552 с.
20. Афанасьев В.Н., Колмановский В.Б., Носов В.Р. Математическая теория конструирования систем управления. – М.: Высшая школа, 2003. – 614 с.
21. Булатов А.В., Кротов В.Ф. О численном решении линейно-квадратичной задачи оптимального управления двойственным методом // Автоматика и телемеханика. – 2009. – № 7. – С. 3-14.
22. Асеев С.М., Бесов К.О., Кряжимский А.В. Задача оптимального управления на бесконечном интервале времени в экономике // Успехи математических наук. – 2012. – Т. 67, вып. 2 (404). – С. 3-64.
23. Мурзабеков З.Н., Мурзабеков А.З. Оптимизация одного класса управляемых нелинейных систем на конечном отрезке времени // Проблемы информатики. ИВМ и МГ СО РАН. – 2012. – № 3. – С. 5-9.
24. Айпанов Ш.А., Мурзабеков З.Н. Аналитическое решение линейно-квадратичной задачи оптимального управления при наличии ограничений на значения управления // Изв. РАН. Теория и системы управления. – 2014. – № 1. – С. 87-94.
25. Квакернаак Х., Сиван Р. Линейные оптимальные системы управления. – М.: Мир, 1977. – 652 с.
26. Кудинов В.А., Карташов Э.М. Гидравлика. – М.: Высшая школа, 2008. – 199 с.