Novosibirsk programming school: a historical overview*

A.G. Marchuk, F.A. Murzin, A.A. Bulyonkova, I.A. Krayneva

Abstract. The paper gives the background and history of the foundation, in 1990, of the Institute of Informatics Systems, Siberian Branch, Russian Academy of Sciences (IIS SB RAS). The institute is rightfully considered the successor to the academic tradition of the Novosibirsk programming school created by Academician A.P. Ershov, head of the Programming Department with the Computing Center, Siberian Branch, USSR Academy of Sciences (1957–1988). Today, Ershov's programming school, whose establishment dates back to the 1960s, continues to develop. In the 1990s, it was transformed into a separate entity, the Institute of Informatics Systems SB RAS. Here, theoretical studies into the topics outlined by A.P. Ershov, his colleagues and disciples are developed and put into practice. The IIS SB RAS is the only research institution in Siberia that focuses on computer science. The genesis of the programming theory and its current development by A.P. Ershov's colleagues and successors have been traced. The IIS SB RAS and Novosibirsk school of programming base their activities on the technological and methodological culture of system engineering, program design and development, which also applies to software engineering. The Novosibirsk programming school as a research and educational phenomenon does not concentrate solely on the good results in theoretical and practical programming. The researchers of the IIS SB RAS actively support social projects on education and informatisation of humanities.

Keywords: science in Siberia, Novosibirsk programming school, software development, Alpha-compiler, Institute of Informatics Systems, A.P. Ershov.

1. Academician Andrei Petrovich Ershov (1931–1988)

The Institute of Informatics Systems was named after Andrei P. Ershov, whose formation as a scientist was influenced by Alexey A. Lyapunov's ideas in automatic programming. Ershov became widely known in this country and abroad from the first steps of his scientific career at the Computing Center of the USSR Academy of Sciences. He is the author of the first Soviet monograph on automatic programming "Programming Program for the BESM Computer" (1958).

In 1957, on the invitation of Academician S.L. Sobolev, A.P. Ershov founded the Programming Department of the Institute of Mathematics and

---

Computing Center, Siberian Branch, USSR Academy of Sciences. The Department ran a number of fundamental projects in the field of system and theoretical programming, which greatly contributed to the foundation and development of programming, to the program schemata theory, parallel programming, program verification, and artificial intelligence.

In Novosibirsk, A.P. Ershov defended his Candidate of Sciences thesis in 1962 and his Doctor of Science thesis in 1966. In 1971, Ershov was elected a Corresponding Member of the USSR Academy of Sciences and in 1984 he became a full member (Academician) of the Academy.

The Novosibirsk team of programmers grew into a regional school of system and theoretical programming. Its distinguishing feature was that all the projects were carried out on an original theoretical and methodological basis. The school pioneered scientific methods of programming team organization and the general principles of programming as a new singular activity with an emphasis on programming technology.

Characteristic of A. P. Ershov's scientific work was the ability to identify the most promising areas of both academic research in computer science and its application to the wide spectrum of daily life needs.

A.P. Ershov introduced the term *informatics* in the context of the fundamental natural science that studies the processes of computer data processing and communication. The observation that the concept of information is cross-discipline and the philosophical consideration of universal cognitive processes as informational lead him to the comprehension of the fundamental nature of informatics. He considered informatics as a science that integrates various aspects of computer science, software development, and their applications.

The Novosibirsk programming team grew considerably in the 1970s. The System Software Design Office and later the Novosibirsk Branch of the Institute of Precise Mechanics and Computing Techniques, USSR Academy of Sciences, were organized. The research area expanded constantly, both following the development of the ideas put forward by the team members and in response to the requests from the industry (mixed computations, artificial intelligence, ES EVM projects, VCKP, etc.). Many young talented students actively participated in those projects.

The academic standing and outstanding personality of A.P. Ershov, head of a department in the Computing Center, made him an informal leader of the national programming community. The numerous conferences in which A.P. Ershov participated became unforgettable events in the life of the community. The Commission on Software Systems of the Academy of Sciences, which was led by A. P. Ershov from 1979 to 1988, became a nation-wide consolidating forum of programmers.

A.P. Ershov made great efforts to establish and deepen both formal and informal relationships with the world scientific community, which made it

possible for Soviet programmers to keep up with the main trends in programming. In particular, contributing to this was his large library, which is still of interest. The library is located at the Institute of System Informatics and has the status of a memorial library.

Ershov became the ideological father of educational informatics by developing the very concept of teaching programming to schoolchildren; he was among the first in the country to realize the importance of personal computers on all levels of the educational paradigm. He played a decisive role in the formation and implementation of the national informatisation programme in 1984-1986.

A.P. Ershov's contribution to the theory and practice of programming was recognized both nation- and world-wide. In 1974, he was appointed Distinguished Fellow of the British Computer Society; in 1981 he received the Silver Core Award for services rendered to IFIP. Ershov was awarded the Soviet state awards: the Order of the Red Banner of Labor (1967, 1975, 1988) and the Order of the Badge of Honor (1981). In 1984, he became the winner of the prestigious A.N. Krylov prize for his work "The theory and application of mixed computation"; in 1985, the winner of the Prize of the USSR Council of Ministers for the development and introduction of advanced technology of automated program design for specialized embedded mini- and micro-computers and supporting mainframe-based adjustable development tools.

In memory of the outstanding scientist, talented teacher, and bright personality, the Institute of Informatics Systems SB RAS holds the international Ershov Informatics Conference ("Perspectives of System Informatics", or PSI); in April, the month when Ershov was born, the annual "Ershov lectures on computer science" are held; the Ershov Prize for young scientists was instituted by SB RAS; annually, Academician A.P. Ershov Summer Schools for Young Programmers are organized. [1]

2. The Programming Department of the Computing Center, USSR Academy of Sciences, as the forerunner of the IIS SB RAS

The Institute of Informatics Systems was set up in 1990. Among its founders and pioneers were colleagues and students of Andrei Ershov from the Programming Department at the Computing Center, SB, USSR AS. Essentially, the IIS inherited both the chief research subjects and the Department style. The previous experience served as a launch pad for future research progress, and nowadays, speaking about the IIS, we never fail to mention the results achieved at the dawn of programming in this country.

As we mentioned, the Programming Department of the Institute of Mathematics together with the Computing Center of the Siberian Branch

of the USSR Academy of Sciences was created in 1957. All applied research done by our team in the various areas of system programming was methodologically based on theoretical models. Below we discuss the most notable academic and applied results of the Programming Department.

2.1. The establishment of the Department: Programming languages and systems

The first widely recognized project of the Programming Department was the ALFA system – an optimizing compiler from the ALFA language, which was an extension of Algol 60 for M-20 computers (1957–1964). The renowned ALFA compiler, considered one of the best in its class, was used for research and engineering tasks at the Computing Center, SB USSR AS, and all over the country. The ALPHA system became the basis for ALGIBR cross-compiler to BESM-6 and ALPHA-71 compiler for M-220.

In 1970, a joint effort of the researchers from the Programming Department, other departments of the CC SB USSR AS and students of Novosibirsk State University was launched to develop ALFA-6 for BESM-6. The ALFA-6 system was used in various Soviet research centers until the last of BESM-6 were dismantled.

In the late 1960s, several innovative languages and systems were created in the Department, such as the SIGMA machine-oriented language, based on the principles of extensibility and tuning to a given object language, and EPSILION programming language, oriented towards symbolic computation. In 1967, an EPSILON compiler appeared for the M-220 type computers, followed by a full EPSILON compiler for BESM-6. The implementation of Lisp, a programming language developed by J. MacCarthy, became an example of international collaboration. The Lisp system for BESM-6 was in use at the CC SB USSR AS practically unchanged until the computer became out-of-date.

In 1971, the BETA project was launched. Initially the intention was to create an open compiling system with a high level of global program optimization, which would encompass the whole class of existing high-level imperative languages, from FORTRAN to Algol 68, and would be oriented towards producing programs for most of the existing computer architecture. In essence, that meant creating a powerful basic compiling system with the common internal language, extensive optimizer and a set of target compilers, and in principle, the task was solved. Simula 67, Pascal, Ada, and Modula-2 compilers were implemented as a result of the BETA project.

2.2. Operating systems: new challenges

Technical progress brought to light new ways of interacting with computers. The AIST project (Automatic Information STation) brought together a wide

range of studies on software and system architecture. As a part of this project and as a result of a joint effort of engineers and programmers, the first Soviet time-sharing system AIST-0 on the basis of M-220 and Minsk-22 was created. The project continued as AIST-1 for the BESM-6 computer. However, works on the project ended in the early 70s, when the DISPAK OS appeared, which embodied the major time-sharing ideas of the time.

In 1976, the works on the VTsKP – a Shared Multi-Access Computing Center – started. The main objective was the networking of high performance computers, such as ES, BESM-6, and Elbrus, for the needs of the Siberian Branch institutes.

2.3. Information systems and databases

In 1967, Andrei Ershov challenged his postgraduate student, A. Zamulin, with a quest to study the feasibility of real-time information search with feedback from the user to the information system. This initiative became the foundation of a new research area of the Department. A program for searching for bibliographical information in the dialogue mode and the Informator program for providing users with information on the AIST-0 components and instructions how to use them were created within the framework of the AIST-0 time-sharing system. It is notable that this program became a prototype for modern help utilities.

The next project was the KADR information system, which was successfully implemented in the personnel departments of various companies. The experience gained from this project allowed to formulate the general guidelines for creating a new general-purpose information search system, VEGA, suitable for a wide range of applications and oriented towards non-programmer users. In time, other systems appeared, but none of them gained wide acceptance due to rapid hardware changes and emergence of numerous new-generation commercial software products.

2.4. Artificial intelligence

In the early 1960s, A.P. Ershov became interested in the problems of natural language communication with computers, and gradually artificial intelligence took a prominent place in the research interests of the Programming Department.

The research group directed by A.S. Narinyani grew into the Artificial Intelligence Laboratory, and quickly took a noticeable, and in some cases a leading position in the booming domestic and international studies in artificial intelligence. By the early 1980s, A.S. Narinyani developed a framework of the underdetermined models technology (now treated as a variety of the later concept of constraint programming), which, together with the semantics-oriented natural language analysis, knowledge-based inference

techniques and intelligent systems construction technology still influences the research conducted by the Artificial Intelligence Laboratory, now based in the IIS SB RAS.

Under the guidance of Yuri A. Zagorulko, the software technological complex Semp-Tec was developed. It was designed to create semantic processors – a type of program systems providing efficient knowledge representation and processing based on semantic webs and production systems. The Semp-Tec technology enables the development of semantic processors that can be used as an intellectual core of expert systems, logical databases, advanced CAD, process control systems, simulators, etc.

In time, the Semp-Tec complex evolved into Semp-TAO, an object-oriented environment, whose knowledge representation model, apart from the classic tools of knowledge representation and processing, included methods of constraint programming.

2.5. The theory of programming

Research in the field of programming theory that started in the Programming Department continued naturally as a part of the research program conducted in the Institute of Informatics Systems; so, speaking about our past results, we acknowledge their further development as well.

The experience of the Department's compiler projects stimulated the development of the theory of program schemata – both sequential and parallel. Work on program schemata focused on the problems of equivalence and transformations of Ianov's program schemata. A.P. Ershov and his disciples found a general criterion of transformation rules locality for Ianov's schemata, proved the logical independence of transformation rules, built a complete system of transformations with identity relation, proved the decidability of logic-thermal equivalence, and so on.

In the theory of parallel programming, classical results belong to V.E. Kotov and A.S. Narinyani. They offered one of the first models of parallel programs – non-deterministic asynchronous automata. This approach was further developed with regard to parallel operator schemata and data flow schemata by V.A. Valkovsky, E.V. Trishina, and I.B. Virbitskaite.

In 1972, A.A. Baehrs proposed a novel computational model based on data flow graph that took into account a variety of data types and had graph folding macro facilities. V.E. Kotov and L.A. Cherkasova created a series of algebraic languages for the description and analysis of parallel non-deterministic processes based on Petri nets.

The team lead by I.B. Virbitskaite developed comparative semantics for various modifications and extensions of Petri nets, as well as algorithms for the verification of real-time net-based models using temporal logics and equivalent transformations.

The IIS Laboratory of Theoretical Programming developed the NetCalc system for analysis and simulation of net-based models. NetCalc is an experimental integrated software system for the design, analysis, and simulation of distributed systems based on various generalizations of Petri nets. It includes customizable graphics editor for hierarchical net structures, analyzer of structural and behavioral properties of models, and a simulation and debugging unit.

The works on mixed computation, initiated by A.P. Ershov, resulted in a wide range of studies covering various aspects of adapting programs to the specific conditions of their use.

V.E. Itkin studied philosophical and mathematical foundations of partial information processing, and together with A.P. Ershov defined and studied the three basic schemes of mixed computation for imperative programs: cut-through, dotted, and polyvariant. M.A. Bulyonkov developed the polyvariant mixed computation scheme that allowed him to build a self-applicable partial evaluator and in 1985 implement all Futamura projections.

B.N. Ostrovsky studied the use of mixed computation for the automatic generation of language specific parsers with the focus on practical application. The advances in the theory and application of mixed computation formed the basis for various analyses and tools for software re-engineering, which were developed in the joint projects of IIS and the Terkom company in 1995–2007 [2].

2.6. Automated publishing systems

When the Russian translation of the book "Revised report on the Algorithmic Language Algol 68", done by A.A. Baehrs, was being prepared for publication, it turned out that the domestic printing houses were unprepared to cope with polygraphically complex text in two languages, with six different typefaces and a mind-boggling layout. There was an urgent need for a modern phototypesetting system.

SAPFIR system for the First Model-Printing House supporting the electronic preparation of publications was developed in 1975–1980. The system was implemented on ES computer with Linotronic 500 imagesetter as an output device.

The RUBIN project for the PRAVDA newspaper was developed in the frames of international cooperation between the USSR and the People's Republic of Poland. The Soviet side was represented by the Pravda publishing office, the Computing Center, SB USSR AS, and M.V. Keldysh Institute of Applied Mathematics, while the MERA-Blone precision mechanics factory and the COBRESPU – Center of television equipment – represented Poland. The project was carried out in 1978-1989 and was aimed at the creation of hardware and software foundation and a system of workspaces for profes-

sional publishers. RUBIN gradually waned because of the well-known events of the late perestroika.

The RUBIN system consisted of a central computational complex with a large information and reference database for the analysis and verification of published materials and forward planning of newspaper issues, and a peripheral local network of workspaces, based on the MRAMOR workstation, for the publishing house employees. In 1987, a pilot batch of 21 MRAMOR workstations with 40 workplaces was produced, and its serial production was prepared. The project was implemented under the auspices of the Council for Mutual Economic Assistance (CMEA) and was discontinued after its collapse.

A.P. Ershov was the scientific adviser of these projects, and A.A. Baehrs was the chief designer [3].

2.7. School informatics

A.P. Ershov considered teaching programming and, more generally, computer science at school as a realization of his thesis "programming is the second literacy". G.A. Zvenigorodsky, Y.A. Pervin, N.A. Yunerman, and other interns and graduate students initiated the Young Programmers School that eventually turned into an all-Union or even international club devoted to working with children interested in computers.

Computer graphics and various concepts of teaching programming languages served as the basis for the curriculum of the Young Programmers School. A group of university and high school students led by Gennady Zvenigorodsky created the graphics system SHPAGA, which was intuitively easy to learn for children, and an integrated system "Shkolnitsa" (Schoolgirl) for the first school computer AGAT. The system included the entry-level programming language ROBIK and the more sophisticated training and production language RAPIRA.

The main functions of the Shkolnitsa system were also implemented for Yamaha PC and Soviet-produced Elektronika-UKNTS. The work was directed by Lydia V. Gorodnyaya and done on the request of the USSR Ministry of Education.

A.P. Ershov and V.M. Monakhov advised the group of authors who wrote the first textbook for high school students and a manual for teachers "Fundamentals of Informatics and Computer Technology". The books were translated into all languages of the Soviet republics.

At present, the work on school informatics is continued by the IIS Group of Educational Informatics.

3. New generation computational tools and systems

3.1. The MARS project

In 1975, the Laboratory of the Theory of Computing Processes led by
V.E. Kotov was formed within the Informatics Department of the Computing Center, SB USSR AS. The project MARS (abbreviation for "Modulnye Asinkhronnye Razvivaemye Sistemy", the Russian equivalent for "Modular Asynchronous Open Systems") became the first major work of the
team. The concept of building next-generation computers was proposed by
G.I. Marchuk and V.E. Kotov in their joint work in 1978. The work laid
down and justified the principles of organizing the computational process.
The analysis conducted was based on pioneering asynchronous communication models, and architecture was viewed as a natural realization of a
computational model.

In 1981, the team led by Yu.L. Vishnevskiy began developing the hardware. Since the existing components could not satisfy the future needs of
the project, A.G. Marchuk started creating VLSI design tools in cooperation
with the Soviet microelectronics industry.

The temporary scientific and technical group START, created by the
USSR State Committee for Science and Technology for advance research and
development in the field of intelligent new generation computers, breathed
new life into the MARS project.

3.2. The Kronos project

Kronos is a generic name of a family of 32-bit processors for building micro-
and mini-computers. Kronos architecture was designed to support high level
programming languages like C, Modula-2, Pascal, Occam, etc., that allowed
implementing the newest ideas in software development and computer use.

The Kronos family of processors was developed within the MARS project
by Kronos research team; its principal developers were Dmitry Kuznetsov,
Alexey Nedorya, Evgeny Tarasov, and Vladimir Filippov. Kronos processors were manufactured as a pilot series mainly for building development
computers designed for software development. The Kronos 2.6 processors
were used for small-lot production of the Kronos-2.6WS workstation. The
workstations were used as development computers in a number of projects
in the USSR defense industry, in particular – for the development of satellite
on-board software at the Scientific and Production Association of Applied
Mechanics (NPO PM, Krasnoyarsk-26). Unfortunately, this development
did not become mass production.

Currently, several Kronos-2.6WS workstations are on display in the Science Museum in London (UK), the Polytechnic Museum in Moscow, Museum of SB RAS, NSU and IIS SB RAS in Novosibirsk.

3.3. The Silicon compiler project

The creation of new generation computers required new hardware components. At that time (1983–1984), the Soviet electronics industry had no future-oriented design tools. Developing each new project always turned into a costly and long process. The idea of the project was to create VLSI design tools, which would generate geometrical drawings, schemes and other information by compiling high-level descriptions of the future device. Abroad, this approach came to be known as "silicon compilation".

The software system for silicon compilation was implemented and tested on real-life problems. For example, one of the microchips from the VLSI chip set for KRONOS was designed using the developed software. These programs were passed on to industry and used in Microprocessor Scientific and Production Association (Kiev), Integral Scientific and Production Association and Scientific Research Institute of Electronic Computer Machines (Minsk), and in Vostok Scientific Research Institute (Novosibirsk).

4. Temporary scientific and technical group START

In 1983, Japan announced a national project aimed at creating fifth-generation computers. The aim was to increase substantially the "intelligence" of the new systems. The Soviet State Committee on Science and Technology, headed at the time by Gury Marchuk, demanded that the Japanese project be evaluated, and a corresponding Soviet project be developed. A group of computer architecture and software specialists was formed, led by Vadim Kotov; members of the team included A.S. Narinyani, A.G. Marchuk, E.H. Tyugu (Tallinn), V.M. Bryabin (Moscow), and others.

The group worked out and published a concept (marked "Confidential") focused around the MARS architecture and intellectualization tools developed by Soviet specialists. The State Committee approved the project and suggested that Vadim Kotov build a creative team to make experimental samples of new-generation computers within a short time (three years).

The temporary research and engineering team START started to work on April, 1, 1985. The budget of START was 12 million rubles for three years and over a million US dollars for purchasing the necessary equipment abroad.

According to the terms of reference, the team was to develop several computers, basic software, instrumental programming and design systems and various intellectualization tools. START fully encompassed the MARS-M supercomputer project. Other projects were centered around the successful Kronos processor.

As a result of the START project in the Artificial Intelligence Laboratory, several new generation intelligent systems appeared: UNICALC, NEMO+,

SEMP-TAO etc.; stemming from these systems were the future joint projects of the IIS AI Laboratory and the Russian Research Institute of Artificial Intelligence, founded by A.S. Narinyani in 1988–1991.

Within the START project timeframe, an impressive set of hardware and software systems and tools was implemented and tested. The list of products include MARS-T parallel system with transputer-like architecture; PIRS workstation (developed by Estonian team); Kronos 2.6 WS Workstation; processing unit based on the Gamma device; Excelsior operating system; compilers for Modula-2, Oberon, and FORTRAN; VLSI design application with elements of silicon compilation; a number of intellectualization tools, BARS programming system for sequential computers. An integrated development environment was created. The main components of this environment were successfully put into operation in the Mechanical Engineering Research Institute (Moscow). More than a hundred developers participated in the START project. For many of them, the START years proved to be the most creative and successful period in their personal academic career.

5. Institute of Informatics Systems

The success of the START project played a significant role in the decision to set up a separate research institute on its basis. The Institute of Informatics Systems (IIS) was founded in accordance with the Order of the USSR Council of Ministers N 1970-r dated 10.11.89, the Order of Council of Ministers of the RSFSR N 1046-r dated 21.11.89, and the Decision of the Presidium of the USSR Academy of Sciences N 161 dated 30.03.90. By the same Decision, DSci Vadim E. Kotov was appointed director of the Institute, and the following research areas were approved:

- theoretical foundations of programming, parallel processing, and artificial intelligence;

- design and architecture of advanced computers, systems and complexes;

- advanced computer system software; expert system and knowledge base software;

- methodology of effective use of data-processing network technology[1].

Later, these areas were specified. In accordance with the Decree of the Presidium of the RAS N 268 of 20.08.97, they are the theoretical and methodological bases for the creation of computer systems, including:

- theoretical foundations of computer science;

---

[1]IIS website – http://www.iis.nsk.su

- methods and tools for reliable and efficient software construction;

- artificial intelligence methods and tools; and

- system and application software for advanced computer systems and networks.

In 1991, D.Sc. Igor V. Pottosin was placed at the head of the IIS. In 1995, the Institute of Informatics Systems was named after Academician A.P. Ershov. In March 1998, DSci. A.G. Marchuk was elected as the IIS director.

Looking forward to the 25th anniversary of the Institute, we are reviewing some of our results and presenting the main areas of theoretical and practical research.

6. System and application software

6.1. Collaboration with the information satellite systems OJSC

In the late 1980s, according to a government decree, Modula-2 was adopted as the base language for the software development of embedded systems. At the same time, Igor V. Pottosin and his team launched the SOKRAT project, which marked the beginning of cooperation with the Krasnoyarsk NPO PM (Scientific and Production Association of Applied Mechanics; now, Academician M.F. Reshetnev Information Satellite Systems OJSC) – the main Russian satellite manufacturer and a leader in space industry.

The SOKRAT system is an experimental development aimed at the creation of integrated tools for building software with higher performance and reliability. It focuses on cross-development tools and is intended for introducing the methods for ensuring efficiency and reliability into the real-life technology of software development for embedded computers. SOKRAT was created as an open, contributed system.

The experience gained from the SOKRAT project was further developed in the frames of the cooperation between the IIS SB RAS and ISS JSC. By 2002, the IIS group headed by A.D. Khapugin fully implemented a cross-system (PCB M2-1750) for the new target platform BTsVM OBC-1750.

Later, the information system ASPID (Archive Maintenance of Software Projects and Documents) was developed for the ISS JSC. It serves to create and maintain an electronic archive for project maintenance of on-board software (OBS) components, and an archive for OBS builds and releases in the development and long-term maintenance of spacecraft on-board software, as well as for the centralized storage and management of all OBS configuration control documents. The ASPID system was approved and put into operation.

6.2. Technology of digital historical factography

In the last decade, a group of researchers led by Alexander G. Marchuk formulated the principles of factual databases designed for storing historical information. They built a non-specific data ontology, designed a general architecture for archive information systems, developed a technology for archival documents and data collection and processing, and implemented interfaces for information viewing, navigation and retrieval. The issues of original documents handling and the problem of reliable storage of their electronic images have been resolved. A number of electronic archive systems were created, such as Academician A.P. Ershov's Archive http://ershov.iis.nsk.su, Siberian Branch of the Russian Academy of Sciences Chronicle http://chronicle.iis.nsk.su, and the Historical portal of NSU Faculty of Mathematics and Mechanics http: //www.globalmmf.ru. The work on filling the Electronic Photo-Archive of SB RAS http://soran1957.ru and the Open SB RAS Archives http://odasib.ru is being continued [4].

7. Theoretical foundations of computer science

7.1. Graph theory application in programming

A.P. Ershov considered graphs to be the fundamental construct in programming having an inexhaustible expressive power. It is impossible to imagine the current state of computer science and programming without the use of graph-theoretic methods.

In the late 1980s, V.N. Kasyanov and V.A. Evstigneev started work on an encyclopedia of graph algorithms for programmers[2]. The authors based the structure of the encyclopedia on the natural division of graph-theoretic algorithms and methods employed in programming into classes defined by the graph types they use. Unlike Donald Knuth, who used the machine-level representation of the fundamental algorithms in his books "The Art of Computer Programming", the encyclopedia authors focused on the high-level description of algorithms in terms of a specially designed pseudo-language (lexicon) containing traditional mathematical and high level programming language constructs [5].

7.2. Sisal-based parallel programs design

The parallel programming language Sisal 3.2 and its compiler were developed as part of an effort to create the design tools for high-quality portable software for parallel computing on inexpensive personal computers.

---

[2]See the IIS staff bibliography at http://www.iis.nsk.su/preprints/papers/2013 (free access)

Sisal 3.2 is a functional programming language which has an implicit parallelism, ensures deterministic results, and contains structures such as loops and arrays. Sisal 3.2 focuses on supporting scientific computing and represents the evolution of Sisal 90 toward the support of advanced inter-module communication, multilingual and object-oriented programming, as well as providing pre-processing and annotated programming features. The methods for optimizing the compilation for Sisal 3.2 have been developed, and an experimental implementation of an optimizing compiler for the .NET platform has been performed. A Sisal runtime library of C# classes has been implemented [6].

7.3. Using colored Petri nets to analyze and verify telecommunication systems

Petri nets are one of the most popular parallel system models. At the IIS, this field has been studied by I.B. Virbitskaite and her disciples. A method of translation of the standard specification language SDL for telecommunications systems into colored Petri nets (CPN) is proposed. A software system STSV (SDL Telecommunications Systems Verifier), which includes a translator from SDL into CPR and a model checking CPR verifier, has been developed and implemented. The STSV system is applied in the study of a well-known problem of feature interaction in telecommunication systems.

An experimental tool SPV (SDL Protocol Verifier) has been designed and implemented for the modeling, analysis and verification of communication protocols in the standard executable specification language SDL. Standard colored Petri nets, as well as modified CPR called hierarchical timed typed nets (HTT-nets) are used for modeling communication protocols [7, 8].

7.4. C-light verification-oriented language

C-light Verification-Oriented Language is a representative subset of C. The important features of C-light are deterministic expression semantics, limited use of **switch** and **goto** statements, and the dynamic memory management using **new** and **delete** C++ operations instead of the standard C library functions. A two-level approach is used for the verification of C-light programs, comprising the steps of the translation of C-light into its kernel – C-kernel language – and the generation of verification conditions using axiomatic semantics of C-kernel. The rules of translation from C-light into C-kernel and the method of a formal proof of their correctness are described.

The project is led by Valery A. Nepomniaschy [10].

7.5. Operational ontological semantics of programming languages

Operational ontological semantics is the new approach to the formal specification of programming languages that describes the programming language ontology and its operational semantics in one context. The method was tested on a number of constructs and mechanisms of modern programming languages, and can be used in the development of the formal specifications of industrial programming languages such as C / C++, C#, and Java [11].

8. Artificial intelligence methods and systems

8.1. Web-based knowledge portal for meaningful access to scientific knowledge and information resources of a given subject area

The concept and implementation of customizable web-based knowledge portal architecture aims at providing meaningful access to systematic knowledge and information resources in a given field of knowledge. The developed information model of the knowledge portal is based on the ontology system including the ontology of scientific activities, ontology of scientific knowledge, and ontology of a particular discipline. The knowledge portal has been customized for the Archaeology and Ethnography subject area and installed and configured on the server of the Institute of Archaeology and Ethnography SB RAS. The project is led by Yuri A. Zagorulko.

8.2. The construction of multilingual thesauri for ontology-based information systems

An approach to the construction of multilingual thesauri for ontology-based information systems has been developed. According to this approach, the thesaurus is built as a linguistic complement of ontology and contains the terms of the problem and subject domains of the information system, by means of which the concepts of ontology are presented in texts and user queries. The approach was tested in the development of a multilingual thesaurus for a knowledge portal on computational linguistics.

8.3. Logical methods for representing the semantics of natural language texts

The work is an effort to systematize machine-oriented logical methods for the analysis of natural language texts and to develop some new algorithms of mapping predicates and formulae of the first-order predicate calculus to the texts in a natural language. Tatiana V. Batura and Feodor A. Murzin examined the possibility of adapting certain structures of mathematical logic

to the study of natural language texts and on this basis gave a new interpretation of the concept "the meaning of the text". They suggested possible applications of these methods to language processing systems that use multi-module memory with parallel access to information.

9. Analysis and processing of large data volumes

9.1. Well logging using radioactive sources

In 2003–2004, on commission from Tyumenpromgeophysika West-Siberian Corporation, a number of algorithms for processing signals generated by radioactive oil well logging were developed and implemented. The resulting software package Spectrum Analyzer provides a wide spectrum of functionality: loading, viewing and processing of the initial amplitude and time spectra; calculation of analytical parameters; calculation of natural radionuclide concentrations; export of the results in the LAS format, used in geophysics. The application area of the package, competitive with world analogues, is oil field operation [9].

9.2. Analysis of information from social networks

The research offers a variety of quantitative characteristics, relations, and sets that can be calculated based on the information obtained from social networks. The data extracted are divided into three categories: user's personal data (such as the name, nickname, time of registration, etc.); user messages and their linguistic characteristics; and user communications. The implemented prototype software package contains the modules performing data extraction, analysis, and visualization. Data extraction module allows us to extract data from the major social networks Twitter, Facebook, and VK. Our research partners are the Institute of Mathematics and Mathematical Modeling MES RK and Suleyman Demirel University (Almaty, Kazakhstan).

9.3. Complex data analysis based on Oracle Business Intelligence technology

The work was performed for one of the largest regional banks in the Urals. Dmitry Semich developed a model, in which the data from the bank's trading day program (CFT IBSO) and from the VISA transactions processing program are put in a database and analyzed hourly. The model was extended with a special business process management module. The project is aimed at the future development of a data-mining system based on Oracle BI technology applied to intelligent control systems. The results of this work are presented in a monograph, in which the author describes his

experience in building a large data repository integrated with a business processes management system.

9.4. Situation analysis and prediction of the Russian transportation network development based on the MIX-PROSTOR software system

The MIX-PROSTOR software system automates research and perspective analysis of the development of the Russian transportation network. It is used to analyze the options of the future development of the Russian transportation network in the context of the global transportation infrastructure. The results can be applied in the economic studies of transport systems and flows. The work is fulfilled in collaboration with the Institute of Economics and Industrial Engineering SB RAS.

9.5. High-precision methods for single-frequency GLONASS / GPS receivers with additional sensors

To detect location, a high precision kinematic method has been developed using additional sensors: inertial MEMS (accelerometer, gyroscope, and magnetometer) and vehicle speed sensor. The maximum number of additional data for error correction is taken into account. To correct the ionospheric delay, the IONEX data on the state of the ionosphere are used. All the data are available on international servers. The method developed allows determining a location with an accuracy better than 1 meter (about 50–70 cm), which is sufficient for some applications, for example, in road construction. In addition, the use of sensors allows for detecting location for a long period of time in low-reception conditions.

9.6. Multi-agent traffic control system

Street lights equipped with special sensors serve as smart devices actively communicating with each other and independently regulating traffic. Due to the active exchange of information on road traffic between traffic lights, the system allows one to build forecasts for urban intersections network for the next 10–20 minutes. Thus the system may solve the problem of reducing the delay time at traffic lights for an average driver and prevent traffic jam.

10. Bioinformatics

10.1. Search and analysis of genetic information

In the framework of international cooperation, an interface for maintaining the databases of Biobase (Wolfenbüttel, Germany; Beverly, USA) was created. The interface tools allow retrieving relevant biological information and

analyzing genetic sequences to identify patterns. It is based on two existing Biobase platforms: Proteome and ExPlain. New algorithms were developed for modeling transcription factor binding sites identified from ChIP-Seq experiments stored in international open databases GEO and ENCODE. The resulting set of alignments, peak positions and binding sites models are available in the GTRD database: http://gtrd.biouml.org/bioumlweb/gtrd.html.

10.2. Algorithms for new generation sequencing data processing

Based on the special-purpose research tools developed at the Institute of Informatics Systems, 1.5 terabytes of data obtained by new generation DNA and RNA sequencing was processed. In 22 human and mouse cell lines and tissues, 5151 long intergenic noncoding regions (vlincRNA) were found, which are known to be markers of pluripotency and cancer. With additional 2.5 terabytes of data obtained by sequencing 833 human cell lines and tissues, possible mechanisms of vlincRNA regulation by transcription factors were established by correlation analysis. In particular, the new genes involved in the indexing process of the pluripotent stem cells were theoretically predicted. 400 Gb of data obtained by the new generation RNA sequencing of wounded mouse skin were processed. The work is conducted in collaboration with a foreign partner (Institute of Saint-Laurent, Providence, USA) [12].

10.3. Tasks and flow control in BioUML Sysytem

The BioUML concept of joint research involves creating a comfortable environment for scientists from around the world to carry out joint modeling and analysis of biological processes, working comfortably with separated data and easily sharing them with colleagues, as well as easily transferring a piece of work from one computer to another. The development of this concept involved an automatic connection of the BioUML system to cloud data stores such as Dropbox and Google Drive. Auxiliary methods based on the MapReduce principle allow algorithm parallelization depending on the number of cores available. A module for BioUML interaction with the statistical software package R was developed. Interaction of BioUML with Galaxy, a platform for biological data analysis, was improved.

The study is carried out together with geneXplain GmbH (Wolfenbüttel, Germany).

10.4. Problem solving on clusters

For a number of data analysis problems of the next generation sequencing (NGS), it is reasonable to use cluster computing. Based on BioUML, a prototype system for managing computing clusters in various configurations

was developed. It provides for the direct transfer of tasks to the nodes by Secure Shell protocol (SSH); support for Sun Grid Engine (SGE) and its analogues by DRMAA protocol; interaction with the Amazon Elastic Cloud Computing (EC2); and the ability to run tasks on kvm virtual machines for better resource sharing.

10.5. OpenWorm: C. elegans nematode simulation

OpenWorm is an international collaboration project aimed at building a simulation model of the most primitive multicellular living organism C. elegans. For several years, the methodology and algorithmic framework focused on high-performance parallel computing in the field of biophysics has been developed. The functionality of the previously created 3D-simulator Sibernetic designed to solve problems in the field of living systems biophysics was extended. Based on the Sibernetic simulator, a new generation computer prototype of C. elegans nematode body was created. The project is coordinated by Stephen Larson (USA, University of California, San Diego) and Andrey Y. Palyanov (IIS SB RAS, Russia).

11. Ershov Informatics Conference (PSI conference series)

In addition to their scientific activities, research workers of the Institute are actively involved in various social and educational projects and promote international cooperation started by A.P. Ershov.

The Institute of Informatics Systems regularly hosts the international Ershov Informatics Conference (also referred to as the PSI ("Perspectives of System Informatics") conference series), dedicated to the memory of Academician A.P. Ershov. The conference is attended by dozens of scientists from abroad. Proceedings are published by Springer in its series *Lectures Notes in Computer Science.*

The Institute has extensive international contacts with scientific organizations and companies working in the field of information technology: Microsoft Research, IBM, Google, Samsung, etc.

Members of the Institute regularly travel abroad to participate in international conferences and joint projects. Seven of the Institute researches are members of various international scientific organizations.

12. Research and teaching

12.1. Training of highly qualified specialists

The Institute is a training base for graduate students and doctoral candidates; it has a Joint Dissertation Council DM 003.032.01 for presenting doctoral and master's theses in specialty 05.13.11 – Mathematical and Software

Support for Computers, Computer Systems and Networks (physical and mathematical sciences). At present, the Council undergoes re-registration with the VAK (Higher Attestation Commission), after which the Council will have a new specialty: 05.13.17 – Theoretical Foundations of Computer Science (physical and mathematical sciences).

The Institute is also the base for the Chair of Programming and Computer Systems of the Mechanics and Mathematics Department (MMD) and Information Technologies Department (ITD) of Novosibirsk State University, where the IIS researchers read a lot of core courses and specialized courses and conduct workshops. Annually, a large number (over a hundred) of senior MMD and ITD students take academic and practical training in the Institute's research laboratories.

12.2. All-Siberian Open Olympiad in Programming

This Olympiad has been organized and conducted jointly with Novosibirsk State University since 2000. Its main purpose is improving the quality of students in the field of information technology, development of knowledge and skills in the key areas of professional activity, improving the quality of recruitment to universities through participation of gifted high-school students in the Olympiad.

The Olympiad is a team competition held in two rounds: the Internet round and the on-site round. From 50 to 60 winning teams of the Internet round are invited to participate in the second round. The teams are from the universities in Russia and CIS countries, half of which are the teams of the universities of Siberia and Russian Far East[3].

12.3.   ICPC International Student Programming Contest

Participation in the All-Siberian Olympiad in Programming is a good rehearsal for the the ACM International Collegiate Programming Contest (ICPC), in which Russian students have participated since 1993, and the team of Novosibirsk State University, since 2000. Coached by the IIS and NSU staff members Tatiana Churina, Tatiana Nesterenko, and Elena Bozhenkova, the NSU team regularly reaches the Contest finals, and in 2007, in Japan, it won silver medals.

12.4. A.P. Ershov Summer School for Young Programmers

Preserving the tradition established by A.P. Ershov and his associates, the IIS SB RAS has been holding for many years the annual Summer School for Young Programmers. The main objectives of the School is educational guidance for talented high school students interested in programming, teaching

---

[3]The Olympiad website is at http://olimpic.nsu.ru/

senior students teamwork skills with the use of modern information technology, developing the practical programming capabilities in junior classes, and providing support for teachers of computer science and programming in schools and universities.

Students are introduced to programming as a production activity, its subject matter, its methodology, creative and technological aspects. They familiarize themselves with the concepts of software product, development process, formulation of the problem and its formalization, rational allocation and scheduling, debugging, execution, documentation, and report writing.

The educational process in the Summer School is organized in the form of training workshops. The main objective of the workshop is to execute the full technological cycle within the scope of the task and to present a report at the final conference. The assistant principal of the Summer School is the IIS researcher Tatiana Tikhonova.

13. Conclusion

The Ershov school of programming, which took shape as a research area in the 1960s, continues to develop today. In the 1990s, it was institutionalized as the Institute of Informatics Systems SB RAS and exists at the departments of Novosibirsk State University, where theoretical studies on the range of problems set by A.P. Ershov and his colleagues I.V. Pottosin, G.I. Kozhukhin, V.E. Itkin, M.M. Bezhanova, G.A. Zvenigorodsky and others are carried out and put into practice.

In the 1990s, despite experiencing, like all other areas of Russian science, the negative impact of destructive social and economic processes, the school of A.P. Ershov kept its research and development facilities, which were successfully implemented in the new millennium. At the core of the Novosibirsk school of programming is the technological and methodological culture of systems engineering, program design and development that goes far beyond the mastery of specific coding tools. Apart from focusing on achieving good results in theoretical and practical programming, researchers of the Novosibirsk school of programming actively support socially significant projects in education and informatisation of humanities [13].

References

[1] Krayneva I.A., Cheremnykh N.A. Put' Programmista [The Path of a Programmer]. – Novosibirsk: Nonparel [Novosibirsk: Publishing House "Nonpareil"]. – 2011.

[2] Institut Sistem Informatiki imeni A.P. Ershova. 20 let. [A.P. Ershov Institute of Informatics Systems. 20th Anniversary] / Ed. by A.A. Bulyonkova. – Novosibirsk, 2010.

[3] Baehrs A.A. The MRAMOR Workstation // Perspectives on Soviet and Russian Computing First IFIP WG 9.7 Conf., SoRuCom 2006, Petrozavodsk, Russia, July 3–7, 2006, Revised Selected Papers / John Impagliazzo and Eduard Proydakov (Eds.). – P. 134–141.

[4] Marchuk A.G., Marchuk P.A. Platforma realizatsii elektronnykh arkhivov dannykh i dokumentov [Implementation Platform of Digital Archives] // Proc. of the RCDL'2012 Conf. – 2012. – P. 332–335.

[5] Kasyanov V.N., Evstigneev V.A. Graph Theory for Programmers. Algorithms for Processing Trees. – Beijing: Kexue Publishing, 2006.

[6] Kasyanov V.N. Sisal 3.2: functional language for scientific parallel programming // Enterprise Information Systems. – 2013. – Vol. 7, No. 2. – P. 227–236.

[7] Andreeva M.V., Virbitskaite I.B. Observational timed equivalences for timed stable event structures // Fundamenta Informaticae. – IOS Press, 2006. – Vol. 72. – P. 1–19.

[8] Tarasyuk I.V. Equivalence relations for modular performance evaluation in dtsPBC // Mathematical Structures in Computer Science. – Cambridge: Cambridge University Press, 2014. – No. 24(1). – P. 78–154.

[9] Murzin F.A., Poplevina N.V., Semich D.F. Algorithms and software for detecting oil reservoirs from nuclear logging data // Optoelectronics, Instrumentation and Data Processing. – 2011 – Vol. 47, Iss. 4. – P. 395–405.

[10] Anureev I.S., Maryasov I.V., Nepomniaschy V.A. Two-level mixed verification method of C-light programs in terms of safety logic // Bulletin NCC. Series: Computer Science. – Novosibirsk, 2012. – IIS Special Iss. 34. – P. 23–42.

[11] Shilov N.V., Akinin A.A., Zubkov A.V., Idrisov R.I. Development of the computer language classification portal // Proc. of Ershov Informatics Conf. PSI Series. – Lect. Notes Comput. Sci. – 2012. – Vol. 7162. – P. 340–348.

[12] St Laurent III G., Shtokalo D., Heydarian M., Palyanov A., Babiy D., Zhou J., Kumar A., Urcuqui-Inchima S. Insights from the HuR-interacting transcriptome: ncRNAs, ubiquitin pathways, and patterns of secondary structure dependent RNA interactions // Molecular Genetics and Genomics. – 2012. – Vol. 287, Iss. 11–12. – P. 867–879.