THE ACADEMY OF SCIENCES
SIBERIAN BRANCH

SOBOLEV INSTITUTE OF MATHEMATICS

INTERNATIONAL CONFERENCE
Mathematical Modeling and High Performance Computing in Bioinformatics, Biomedicine and Biotechnology, MM-HPC-BBB-2014

Abstracts

MM-HPC-BBB-2014
Novosibirsk, RUSSIA
June 24–25, 2014

Publishing House SB RAS

Novosibirsk
2014

PROGRAM COMMITTEE

Chairs of program committee:
Kabanikhin S.I., Prof., RAS corresponding member, Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
Kolchanov N.A., Prof., RAS member, Institute of Cytology and Genetics SB RAS, Russia

Vice-chairs of program committee:
Shishlenin M.A., Ph.D., Institute of Mathematics SB RAS, Russia
Chernykh I.G., Ph.D., Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia

Program commitee:
Björn Sommer, Dr., Bielefeld University, Germany
Dehmer M., Prof., Institute for Bioinformatics and Translational Research, UMIT - The Health and Life Sciences University, Austria
Milanesi L., Prof., ITBA, Milan, Italy
Rzhetsky A., Prof., Institute for Genomics and Systems Biology at the University of Chicago, USA
Zola Jaroslaw, Prof., Rutgers University, USA
Aulchenko Yu. , Ph.D. , Institute of Cytology and Genetics SB RAS, Russia
Bektemesov M.A., Prof., KazNU, Kazakhstan
Bocharov G.A., Prof., Institute of Computational Mathematics RAS, Russia
Vesnin A.Y., RAS corresponding member, Institute of Mathematics SB RAS, Russia
Vityaev E.E., Prof., Institute of Mathematics SB RAS, Russia
Gainova I.A., Ph.D., Institute of Mathematics SB RAS, Russia
Glinsky B.M., Prof., Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
Golubyatnikov V.P., Prof., Institute of Mathematics SB RAS, Russia
Golushko S.K., Prof., Design Technological Institute of Digital Techniques SB RAS, Russia
Goncharov S.S., RAS corresponding member, Prof., Institute of Mathematics SB RAS, Russia
Demidenko G.V., Prof., Institute of Mathematics SB RAS, Russia
Erohin G.N., Prof., Kant's Baltic Federal University, Russia
Zagoruiko N.G., Prof., Institute of Mathematics SB RAS, Russia
Zybarev Yu.M., Ph.D., Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
Ilyin A.I., Prof., JSC "NTSPP", Kazakhstan
Koptiug I.V., Prof., International Tomography Center SB RAS, Russia
Kulikov I.M., Ph.D., Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
Kuramshina G.M., Prof., Moscow State University, Russia
Lavrentiev M.M., Prof., Novosibirsk State University, Russia
Likhoshvai V.A., Ph.D., Institute of Cytology and Genetics SB RAS, Russia
Marchuk A.G., Prof., A.P. Ershov Institute of Informatics Systems SB RAS, Russia
Moshkin M.P., Prof., Institute of Cytology and Genetics SB RAS, Russia
Nurseitov D.B., Ph.D., KazNTU, Kazakhstan
Pertsev N.V., Prof., Omsk branch Institute of Mathematics SB RAS, Russia
Romanyukha A.A., Prof., Institute of Computational Mathematics RAS, Russia
Samsonov A.M., Ph.D., St.Petersburg State Polytechnical University, Russia
Snytnikov V.N., Ph.D., Boreskov Institute of Catalysis SB RAS, Russia
Fadeev S.I., Prof., Institute of Mathematics SB RAS, Russia
Fedoruk M.P., Prof., Novosibirsk State University, Russia
Fedotov A.M., RAS corresponding member, Novosibirsk State University, Russia
Chupakhin A.P., Prof., Lavrentyev Institute of Hydrodynamics SB RAS
Shokin Y.I., RAS member, Institute of Computational Technologies SB RAS, Russia
Yagola A.G., Prof., Moscow State University, Russia
Podkolodnyy N.L., Institute of Cytology and Genetics SB RAS, Russia

LOCAL ORGANIZING COMMITTEE

Chair:
RAS corresponding member S.I. Kabanikhin, Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia,

Vice chair:
M.A. Shishlenin, Sobolev Institute of Mathematics, Novosibirsk, Russia,
A.V. Penenko, ICM&MG SB RAS, Novosibirsk, Russia
N.L. Podkolodnye, ICM&MG SB RAS, Novosibirsk, Russia

Scientific secretary
O.I. Krivorot'ko, ICM&MG SB RAS, Novosibirsk, Russia

Organizing committee:
D.A. Voronov, ICM&MG SB RAS, Novosibirsk, Russia
N.S. Novikov, ICM&MG SB RAS, Novosibirsk, Russia
I.G. Chernykh, ICM&MG SB RAS, Novosibirsk, Russia
I.M. Kulikov, ICM&MG SB RAS, Novosibirsk, Russia

ISBN 978-5-7692-1372-4 (т. 2)

978-5-7692-1370-0

Organizers

Institute of Internal Medicine SB RAMS

Sponsors

Russian Foundation for Basic Research

SILVER SPONSORS

Helicon

Limited Liability Company Eppendorf Russia

Genotek

Bio-Rad

intel

InterLabService

BRONZE SPONSORS

Roshe diagnostics

CONTENTS

APPLICATION OF CELLULAR AUTOMATA FOR INVESTIGATION OF POLLUTION INFLUENCE ON MACROHECTOPUS AND COMEPHORUS POPULATION IN THE LAKE BAIKAL

Afanasyev I.V.
Institute of Computational Mathematics and Mathematical Geophysics SB RAS
e-mail: ivafanas@gmail.com

Key words: cellular automata, prey-predator systems, parallel composition, population dynamics, lake Baikal

Motivation and Aim: Baikal is the biggest reservoir of sweet water. Investigation of biochemical processes is required for preserving of the lake. Since anthropogenic influence is usually local, models of biochemical processes must take into account spatial distribution of parameters. Comephorus baikalensis and Comephorus dybowski take leadership in the biomass of fishes of Baikal. Its basic food is Macrohectopus and its own young. The aim of the work is the development of a model of Macrohectopus, Comephorus baikalensis and Comephorus dybowski population dynamics, which takes into account spatial individuals distribution and effect of possible local pollution.

Methods and Algorithms: Each kind of organisms is divided into age groups. There is eight groups of organisms totally. Demographic and prey-predator interactions are defined between groups of organisms [1]. Cellular automata approach is used to build the model. The model proposed is a parallel composition [2] of eight cellular automata; each of them being entitled to model population of organism group. The model takes into account seasonality, organisms' movement, water streams and pollution.

Results: Model verification is done. Assessments of model results by verification parameters differ from assessments given in literature in about 20%. Required software is developed for computational experiments. The algorithms were implemented and parallelized for systems with shared memory. The process under simulation tends to annual oscillations and non-uniform individuals distribution. Annual oscillation is the result of seasonality. Non-uniform distribution is the consequence of water streams influence.

Conclusion: Simulation shower that local pollution influence doesn't spread out of polluted territory. Assessments are presented for the critical pollution leading to total death of individuals, and that of acceptable one, which influence is not observable.

Availability: Program code is available here: https://github.com/ivafanas/ca_baikal

References:

1. V.I. Zorkalcev, I.V. Mokry, V.I. Kazazaeva (2008) Модель взаимодействия трёх пелагических видов организмов озера Байкал, Modern technologies. System analysis. Modeling. 1: 182–193.
2. O.L. Bandman (2010) Cellular Automata Composition techniques for Spatial Dynamics Simulation, Simulating Complex Systems by Cellular Automata. Understanding Complex Systems. 0: 81–115

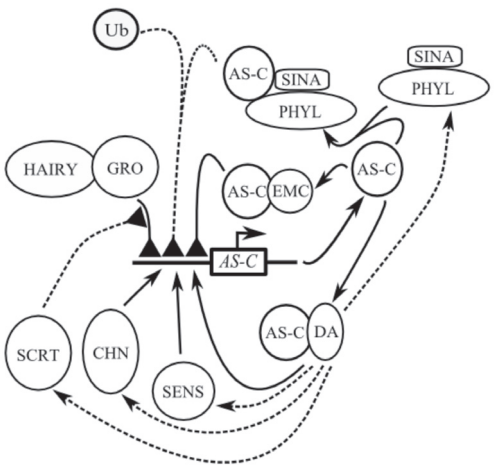
AN EXTENDED MODEL OF *D. MELANOGASTER* MACROCHAETE MORPHOGENESIS GENE NETWORK

Akinshin A.A.[1, 4], Bukharina T.A.[2], Furman D.P.[2, 3], Golubyatnikov V.P.[3, 4]*
[1]*Altai State Technical University, Barnaul, Russia;*
[2]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[3]*Novosibirsk State University, Novosibirsk, Russia;*
[4]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia*
e-mail: glbtn@math.nsc.ru
**Corresponding author*

Key words: drosophila, achaete–scute complex, gene networks mode, phyllopod

Motivation and Aim: The goal of our work was to construct an extended mathematical model describing the changes in AS-C protein content in this cell under the control of CRC. *Methods and Algorithms:* Our approach to modeling of the Central Regulatory Contour (CRC) is based on methods elaborated in [1,2].

Results: We study the model of the gene networks, see the Fig.1, represented by a nonlinear kinetics dynamical system. The AS-C proteins degrade through an ubiquitin-dependent pathway containing the protein PHYL, playing the adaptor's role, and the ubiquitin ligase seven in absentia (SINA). The lifespan of proneural proteins is determined by the protein PHYL. The PHYL accumulation depends on AS-C proteins, which, within the heterodimers AS-C/Daughterless, initiate transcription of the gene *phyllopod*. This self-regulatory loop provides for a balance in the contents of AS-C and PHYL proteins.



Negative feedbacks
GRO ···◄ (AS-C),
HAIRT-GRO ···◄ (AS-C)
etc.

Positive feedbacks:
CHN → (AS-C),
DA-(AS-C)→SINA-PHYL
etc.

Similar simplified model of this gene network without Ub, SINA and PHYL was studied in [1].

Fig.1. The CRC of the *AS-C* expression control.

Conclusion: We find conditions of existence of a single stable equilibrium point of this model, and describe its qualitative behavior.

Supported by SB RAS, RFBR, grants 80, 136, 12-01-00074 and by grant VI.61.1.2 of RAS.

References:

1. T.A. Bukharina et al. (2012) Model investigation of central regulatory contour of gene net of *D.melanogaster* machrohaete morphogenesis. Russ. J. Develop. Biol. **43(1)**: 397-412.
2. V.P. Golubyatnikov, I.V. Golubyatnikov (2011) On multidimensional models of gene networks functioning. Journ. of Applied and industrial mathematics, **5(3)**: 343-347.

MODELING OF TWO-CELLS COMPLEX IN MORPHOGENESIS OF *D. MELANOGASTER* MECHANORECEPTORS

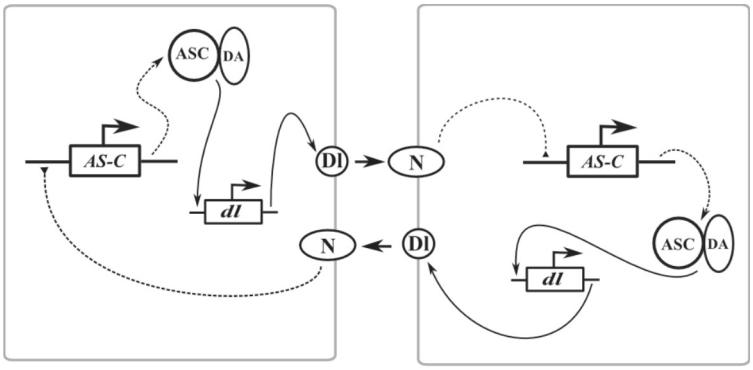
Akinshin A.A.[1, 4], Bukharina T.A.[2], Furman D.P.[2, 3], Golubyatnikov V.P.[3, 4]*
[1]*Altai State Technical University, Barnaul, Russia;*
[2]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[3]*Novosibirsk State University, Novosibirsk, Russia;*
[4]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia*
e-mail: glbtn@math.nsc.ru
**Corresponding author*

Key words: drosophila, achaete–scute complex, gene networks, mathematical model

Motivation and Aim: The morphogenesis of macrochaetes is a multistage process. We consider the first stratification period of the proneural cluster of the wing imaginal disc.

Methods and Algorithms: Our description of phase portraits of Central Regulatory Contour models is based on methods elaborated in [1, 2].

Results: We consider symmetric gene networks model represented by nonlinear chemical kinetics dynamical system (1).



$$\frac{d x_i}{d t} = f(z_i) - x_i;$$
$$\frac{d y_i}{d t} = \sigma(x_i) - y_i;$$
$$i=1, 2; \qquad (1)$$
$$\frac{d z_1}{d t} = \sigma_*(y_2) - z_1;$$
$$\frac{d z_2}{d t} = \sigma_*(y_1) - z_2.$$

Fig.1. The scheme of interaction of two adjacent identical cells K_1 and K_2 in the proneural cluster.

The variables $x_i(t)$, $y_i(t)$, and $z_i(t)$, i=1,2 denote concentrations of the proteins AS-C, Delta, and, respectively, Notch in the cells K_i. The monotonically increasing functions $\sigma(x)$, $\sigma_*(x)$ describes positive feedbacks (AS-C) → Dl and, respectively, positive intercellular feedbacks Dl → N from the cell K_1 to K_2 and vice versa, see the Fig. 1. The monotonically decreasing function f describe the negative feedbacks N ⋯◀ (AS-C).

Conclusion: We find conditions of existence of 3 equilibrium points of the system (1). Two of these points are stable and correspond to two possible variants of the final state of this two cells complex in the proneural cluster. The third equilibrium point P_3 is unstable. We find conditions of existence of an unstable periodic trajectory of the system (1) near P_3.

Acknowledgements: Supported by SB RAS, RFBR, grants 80, 136, 12-01-00074, and by grant VI.61.1.2 of RAS.

References:

1. Yu.A.Gaidov, V.P.Golubyatnikov (2011) On the existence and stability of cycles in gene networks models with variable feedbacks, *Contemporary mathematics*, **553:** 61-74.
2. T.A.Bukharina et al. (2012) Model investigation of central regulatory contour of gene net of *D.melanogaster* machrohaete morphogenesis. Russ. J. Develop. Biol. **43(1)**: 397-412.

USING NOVEL GENERIC STRING KERNEL TO BUILD PAN-SPECIFIC MHC CLASS I PEPTIDE BINDING PREDICTION TOOL

Antonets D.V.[1]*, Grudin D.S.[2]
[1]*State Research Center of Virology and Biotechnology «Vector», Koltsovo, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia*
e-mail: antonec@yandex.ru
**Corresponding author*

Motivation and Aim: CD8+ T-cell epitopes play crucial role in antiviral and anticancer immunity. Reliable prediction of T-cell epitopes remains one of the most important tasks of immunoinformatics since accurate *in silico* identification of potent epitopes could drastically reduce materials and time consumption in comparison to traditional experimental approaches of epitope discovery. The main goal of this work was the development of new statistical models for predicting peptide binding to different allomorphs of MHC class I molecules using novel GS (Generic String) kernel function [1]. The main advantage of GS kernel is its unique ability to consider both substring position uncertainty and physicochemical properties of amino acids, making GS able to compare peptides of different lengths. MHC molecules are extremely polymorphic while the number of MHC variants with sufficient experimental peptide-binding data is very limited. Thus there is a strong interest to develop new pan-MHC specific T-cell epitope prediction tools. Here we applied GS kernel to measure the similarity between different MHC molecules, encoded as pseudosequences of their peptide-binding sites.

Methods and Algorithms: GS, MHC2SK and SupCK kernels were implemented as it was described by the authors [1-3]. To optimize computational performance they were implemented using C++ and integrated into R via Rcpp package. Predictive models were built using support vector regression and relevance vector regression algorithms implemented in R package kernlab. Performance of produced models was assessed using ROCR. Amino acid residues were encoded with PMBEC amino acid similarity matrix [4]. Current predictive models were built using only HLA class I binding data for nonameric peptides only (taken from IEDB; http://immuneepitope.org).

Results and Conclusion: Almost all the models demonstrated good quality of predictions with AUC values ranging from 0.84 to 1.0, and Pearson's correlation coefficients between predicted and experimentally determined pIC50 values of MHC:peptide binding ranging from 0.67 to 0.97. Predictive models built with relevance vector machine approach were significantly less complex as those built using support vector regression. Both GS and MHC-2SK kernels were found to be almost equally effective, and they were found to be superior as compared to SupCK. Next we plan to develop the models using both human and non-human peptide-MHC binding data. Detailed testing results and ROC curves will soon be available at http://tepredict.sourceforge.net/GSK-SVR/.

Availability: http://tepredict.sourceforge.net.

Acknowledgements: This work was supported by research grant from Novosibirsk Region Government.

References:

1. Giguere S. et al. (2013) *BMC Bioinformatics*, **14**:82.
2. Guo L. et al. (2013) *BMC Genomics*, **14**(Suppl 5):S11.
3. Hoffman B. et al. (2010) *BMC Bioinformatics*, **11**:99.
4. Kim Y. et al. (2009) *BMC bioinformatics*, **10**:394.

ON THE APPLICATION OF EXCRETION DATA AS A CRITERIA OF CHOICE BETWEEN MULTIPLE SOLUTIONS OF INVERSE PROBLEM IN PHARMACOKINETICS

Asmanova N.*, Koloskov G., Ilyin A.I.
Scientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan
e-mail: asmanova@inbox.ru
** Corresponding author*

Key words: compartment models, first-order absorption, excretion, nonuniqueness

Motivation and Aim: Among the published papers dealing with the multiple solutions of the inverse problem in pharmacokinetics it is not rare to find a remark suggesting that the excretion data would be helpful to resolve the nonuniqueness of the solution. The purpose of this study is to determine the possibility of usage of excretion data as a criteria of choice between multiple solutions of inverse problem in pharmacokinetics in case of pharmacokinetic (PK) models with first-order absorption and elimination.

Methods and Algorithms: The fundamental difference of the drug pharmacokinetic analysis in plasma/blood and urine is that the amount of drug in plasma can be only measured in terms of concentration C(t), while in case of urine it is the mass of the drug m(t). Analysis of the solution-dependent parameters (rate constant of excretion k_e, rate constant of elimination k_{10} and volume of distribution V_1) as well as solution-invariant parameters suggests that eliminating curves C_{10} (t) corresponding to multiple of solutions [1,2], coincide when expressed in terms of m_{10}(t).

Results: Even if excretion is the only process of drug elimination, the excretion data can not be used as a criteria of choice between multiple solutions, since the pharmacokinetic of excretion is also a part of the problem.

Conclusion: It is concluded that the comparison of the pharmacokinetics of plasma and urine, can be useful only to test the consistency of the model.

References:

1. Asmanova N., Koloskov G., Ilin A. (2013) Coupled solutions of one- and two-compartment pharmacokinetic models with first-order absorption. J. Pharmacokinet. Pharmacodyn., 40: 229–241, DOI:10.1007/s10928-013-9312-6.
2. Koloskov G., Asmanova N., Ilin A. (2014) Application of a biexponential form of two-compartment pharmacokinetic models with first-order absorption in pharmaco-kinetic analysis. International Conference “Mathematical Modeling and High-Performance Computing in Bioinformatics, Biomedicine and Biotechnology” Novosibirsk, Russia, 24–27 June 2014.

ON GEOMETRY OF PHASE PORTRAITS OF SOME LOW-DIMENSIONAL GENE NETWORK MODELS

Ayupova N.B.[1,2], Golubyatnikov V.P. * [1,2]
[1]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;*
[2] *Novosibirsk State University, Novosibirsk, Russia*
e-mail: glbtn@math.nsc.ru
**Corresponding author*

Key words: gene network, mathematical model, cycles, integral submanifold

Motivation and Aim: Multistablity of gene network models is very important both for mathematical modeling of biochemical kinetics, and for its biological applications.

Methods and Algorithms: Our description of phase portraits of gene network models is based on geometrical methods elaborated in [1].

Results: We consider gene networks models represented by n-dimensional dissipative dynamical systems, n=3 or n=4, of the type:

$$\frac{d\,x_i}{d\,t} = f_i(x_{i-1}) - x_i;\ i = 1,..,n;\ i-1 = n \quad \text{for} \quad i = 1, \tag{1}$$

where $f_i(x)$ are threshold decreasing functions, which represent the negative feedbacks in the model: $f_i(x)=A_i=const>2$ for $0\le x<1$, and $f_i(x)\equiv 0$ for $x \ge 1$; i = 1,…n. For arbitrary dimensions n, trajectories of the system (1) are piece-wise linear, their segments are described by the projective transformations, and all these trajectories enter eventually the block $[0,A_1]\times[0,A_2]\times...\times[0,A_n]$ and do not leave it when $t \to \infty$.

For ***n*=3**, we show that the chemical kinetics system (1) has a unique cycle C_3, and we construct piece-wise linear surface (integral submanifold of the phase portrait) which is invariant with respect to the shifts along trajectories of the system (1) and contains C_3. This cycle is stable for sufficiently large values of the parameters A_i, i=1, 2, 3.

For ***n*=4**, we consider **symmetric** systems (1), where $A_1 = A_2 = A_3 = A_4 = A$. As it was shown in [2], for $A > 2$ the system (1) has two stable equilibrium points $S_1 = (A,0, A,0)$ and $S_2 = (0, A,0, A)$ in R^4, and a unique cycle C_4. As in the 3D case above, this cycle is contained in 2-dimensional invariant surface $M^2 \subset R^4$. Now we have done an explicit description of 3-dimensional piece-wise linear integral submanifold M^3 which contains M^2, separates attraction basins of the equilibrium points S_1 and S_2, and is invariant with respect to the cyclic permutation of the variables $x_1 \to x_2 \to x_3 \to x_4 \to x_1$. The cycle C_4 is stable within the surface M^2 and is unstable in the phase portrait of the system (1).

Conclusion: In the cases of higher-dimensional dynamical systems of the type (1), similar considerations can be done as well. Some of them were realized in [2] for n=5.

Acknowledgements: Supported by SB RAS, Interdisciplinary Project 80, and by RFBR, grant 12-01-00074.

References:

1. V.A. Likhoshvai, V.P. Golubyatnikov et al. (2008) Theory of gene networks, In: *System computerized biology*, (N.A. Kolchanov, S.S. Goncharov), Novosibirsk, SB RAS, 397-480.
2. A.A. Akinshin, V.P. Golubyatnikov, I.V. Golubyatnikov (2013) On some multidimensio-nal models of gene network functioning. Sib. Journ. of Industrial mathematics. **16(3):** 3-9.

DYNAMIC INSTABILITIES OF MICROTUBULES

Barlukova A.[1]*, Honoré S.[2], Hubert F.[1], Petit M.[2]
[1]Aix-Marseille University, I2M CNRS UMR 7353, Marseille, France;
[2]Aix-Marseille University, Inserm, CRO2 UMR 911, Marseille, France
e-mail: ayuna.barlukova@etu.univ-amu.fr
**Corresponding author*

Key words: cancer, microtubules, dynamic instability, MT targeting drugs

Motivation and Aim: Microtubules (MTs) are long tube polymers of tubulin, found throughout the cytoplasm. MTs are very important in all crucial cellular processes in cancer progression such as cell division and migration. The aim of the work is to develop new mathematical models that account for the effects of MT targeting agents on MT instabilities.

Methods and Algorithms: We propose a new deterministic mathematical model inspired by the work of P. Hinow et al. [1] to simulate the behavior of a MT population. The model couples transport equations with ordinary differential equations (ODE) with nonlocal terms endowed with suitable boundary conditions for both catastrophe and rescue. The mathematical model is built from biological observations obtained by the pharmacologist of our interdisciplinary research group [2]. Numerical results are obtained in MATLAB by using upwind scheme with adaptive time step for the partial differential equations and the explicit Euler method for the ODE.

Results: We obtain graphs for time evolution of the average total length of MTs in polymerization state and their caps and average length of MTs in depolymerization state (similar to data obtained by kymograph); concentrations of free GTP and GDP tubulin, total quantities of tubulin incorporated in MTs in polymerization and depolymerization states; time evolution of distribution of MTs in polymerization state.

Conclusion: Computational simulations describe diverse concepts of behavior of MT populations with and without impact of drugs. New model allows us to analyze the pharmacological action of different anti-microtubule drugs, including that influence on MT "aging", on MT instabilities. Numerical results are in a good agreement with biological observations.

Acknowledgements: The program is funded thanks to the support of the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the «Investissements d'Avenir» French Government program, managed by the French National Research Agency (ANR).

References:

1. P. Hinow, V. Rezania, J.A. Tuzszynski. Continuous model for microtubule dynamics with catastrophe, rescue, and nucleation processes. Phys Rev E Stat Nonlin Soft Matter Phys. 2009 Sep; 80(3 Pt 1): 031904.
2. R. Mohana, E.A. Katrukhaa, H. Doodhia, I. Smalb, E. Meijeringb, L.C. Kapiteina, M.O. Steinmetzc, A. Akhmanova. End-binding proteins sensitize microtubules to the action of microtubule-targeting agents. Proc. Natl. Acad. Sci. U.S.A. V. 110, 1091-6490, 2013.

APPLICATION OF THE METHODS OF PERSISTENT HOMOLOGY TO CLINICAL DATA ANALYSIS

Bazaikin Ya.V.[1], Chupakhin A.P.[2, 3], Cherevko A.A.[2, 3], Khe A.K.[2, 3]*
[1]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;*
[2]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia;*
[3]*Novosibirsk State University, Novosibirsk, Russia*
e-mail: alekhe@hydro.nsc.ru
**Corresponding author*

Key words: time series analysis, clinical data, persistent homology, barcodes, persistent diagrams

Motivation and Aim: A problem of calculation of the topological characteristics of geometric objects of complex shape arises in various applications. This led to the emergence of computational topology that is developing lately. One of the most common approaches to calculating homology groups is based on the Morse theory, which should be modified when studying discretely defined geometric objects. In this work, persistent homology methods are applied to the data of medical examinations and clinical data (hemodynamic parameters) obtained during operations on the cerebral vessels in the Meshalkin Novosibirsk Research Institute of Circulation Pathology.

Methods and Algorithms: Data analysis of medical examinations for assessment of the endothelial function was conducted using barcode construction. The apparatus reads the data with two independent sensors placed on the fingers of the left and right hands. A wavelet transform is applied to the data obtained, resulting in a two-dimensional array of coefficients. Coefficients are calculated on the "time—scale" plane. For obtained wavelet coefficients, which are discrete functions of two variables, barcodes and their length distribution are calculated.

Using persistent diagrams, clinical data of endovascular examinations carried out during minimally invasive surgical procedures in the Meshalkin Novosibirsk Research Institute of Circulation Pathology, are analyzed. A measuring system with intravascular sensor of the diameter of 0.36 mm was used for measuring the pressure and flow rate in the vessels of the brain. For the data obtained, as for one-dimensional functions, persistent diagram are constructed.

Results: The results obtained show a qualitative difference in barcodes and their distributions for different patients. The barcodes consist of several groups which vary for different study cases. The distribution functions are distinguished by their extrema.

The persistent diagram for intravascular pressure shows two areas of point crowding in addition to a region adjacent to the diagonal. These areas are responsible for the lower-frequency oscillations, which may be related to physiological processes such as breathing. The diagram for velocity has more blurry crowding region, which may be due to the characteristics of the measuring equipment (increased noise level).

Conclusion: The results obtained in this work show the possibility of application of the persistent homology methods to clinical data analysis, which can improve the diagnosing methods.

Acknowledgements: This work was supported by Presidium of the Siberian Branch of the Russian Academy of Sciences (Integration Project 44), the Department of Energetics, Engineering, Mechanics and Control Processes of the Russian Academy of Sciences (Program 2.13.4), Russian Fund for Basic Research (Project 14-01-00036).

COMPUTER MATHEMATICAL AND BIOCHEMICAL MODELING AND SIMULATION OF THE LIFE PROCESSES IN HUMAN ORGANS

Bedelbayev A.A.
Institute of Mathematics and Mechanics KazNU, Almaty, Kazakhstan
e-mail: agyn08@gmail.com

Key words: modelling, biological, chemical processes, human organs, computer

Motivation and Aim: The purpose of research is computer simulation of biochemical processes in human organs in his daily life, based on mathematical and biochemical methods for describing such processes and diverse library of templates conditions change their course based on various scenarios of the changing environment and the nature of their various influences. The results will be invaluable help in the study of diseases of the human body, and the development of the definition of irreversible biological and physical changes and allow it to plan scientifically based methods for its treatment.

Methods and Algorithms: To successfully solve this problem used the combination of biochemical methods and mathematical study of the various processes occurring in the human body in a variety of scenarios surrounding this change in body temperature and conditions of its functioning. Visualized representation for subtle physical and biochemical processes using the latest high-tech and effective information technology, allowing in some cases have not only a supporting role, but also research.

Results: Building representative strictly structured database characteristics specific organ functioning person with deep classification distinguishes the characteristics of his life. Building an effective database of various scenarios of changes in the isolated human body with a broad spectrum of influence on it of the changing environment , which gives the opportunity to quickly get the most comprehensive information in a different format requested. Building an effective technology dispatching dynamic changes of biochemical processes in the human body selected under different conditions by selected characteristics of these processes. Construction of specialized structured in different categories of the characteristic changes in the physical properties of the isolated human organ database at different scenarios of changes in its environment. Construction of mathematical and biochemical algorithms of modeling the functioning of the selected organ of human body under the influence of changing environmental conditions on the various scenarios it changes. Creating a rendered model of the functioning of human body with various characteristics and environmental change scenarios of biochemical processes. Creating an effective acupressure techniques on the course of biochemical and biomechanical processes in isolated human body under the influence of special biochemical preparations. Test a computer model of human body functioning under natural environmental conditions, corresponding to the most typical complaints of patients. Developing standards for treatment of human body by simulating the behavior of selected human body under different scenarios of targeting stroke occurring biochemical processes by introducing different biochemical preparations corresponding to the different methods of treating a disease of the body of man.

CENTROID-BASED ENSEMBLE OF K-MEANS FOR BIG DATA CLUSTERING

Berikov V.B.
Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia
e-mail: berikov@math.nsc.ru

Key words: clustering ensemble, k-means, centroids

Motivation and Aim: Ensemble clustering is aimed to get a stable solution to the clustering problem by combining partial clustering partitions [1]. There are a number of different approaches in the theory and methods of ensemble clustering: median partitioning, co-occurrence matrix averaging, graph-based partitioning, spectral clustering etc. For big data such as large-scale genomic data sets or hyperspectral images, the time and storage complexity of existing ensemble algorithms become intolerably large due to the label correspondence problem or because of the need to compare data objects in a pairwise manner. This paper proposes an efficient approach to the ensemble clustering needn't perform label analysis or pairwise comparisons.

Methods and Algorithms: *K*-means is a popular clustering algorithm due to its simplicity and effectiveness. Conventional methods of collective clustering which use *k*-means algorithms as base elements of the ensemble form a consensus partition by the analysis of cluster labels assigned to data objects. This procedure has a number of drawbacks because of the arbitrary nature of the assignment. At the same time, the centroids produced by *k*-means algorithms serve as cluster prototypes and can be used for the design of the ensemble solution by simple averaging of their coordinates. The algorithm of centroid-based *k*-means ensemble clustering consists of the following steps. **Input**: data set $\mathbf{x} = (x_i)$, $i = 1,..,N$, where $x_i \in R^d$ is a vector of feature values. **Output**: clustering partition of $\mathbf{x}$ on the given number K of clusters.

1. Produce L variants of *k*-means clustering by random initialization of algorithm's parameters (such as initial centroids coordinates, number of centroids etc.).

2. For each i-th object, find closest centroid $C_{i,l}$ in each l-th variant of partition and set $\tilde{x}_i := 1/L \sum C_{i,l}$ (the object is mapped to the average of its prototypes).

4. Apply *k*-means with the given number of clusters K to the amended data set $\tilde{\mathbf{x}}$.

5. Form a final partition of $\mathbf{x}$ by matching cluster labels of $\tilde{\mathbf{x}}$.

Results: The suggested method is theoretically substantiated with the Central Limit Theorem. The computationally effective algorithms for microarray data clustering and segmentation of hyperspectral images are developed on the basis of the method.

Conclusion: The proposed method allows getting efficient solution to the problem of big data clustering.

References:

1. Berikov V. Weighted ensemble of algorithms for complex data clustering // Pattern Recognition Letters. 2014. Vol. 38. P. 99-106.

A MACHINE LEARNING ANALYSIS OF URINE PROTEOMICS IN SPACE-FLIGHT SIMULATIONS

Binder H.[1], Wirth H.[1], Arakelyan A.[2], Lembcke K.[1], Tiys E.S.[3], Ivanishenko V.A.[3], Kolchanov N.A.[3], Kononikhin A.[4], Popov I.[5], Nikolaev E.N.[4, 5], Pastushkova L.[6], Larina I.M.[6]

[1]*Interdisciplinary Centre for Bioinformatics, Universität Leipzig, Leipzig, Germany;*
[2]*Institute of Molecular Biology NAS RA; Yerevan, Armenia;*
[3]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[4]*Talrose Institute for Energy Problems of Chemical Physics, RAS, Moscow, Russia;*
[5]*Emanuel Institute for Biochemical Physics, RAS, Moscow, Russia;*
[6]*Institute of Biomedical Problems – Russian Federation State Scientific Research Center RAS, Moscow, Russia*
e-mai: binder@izbi.uni-leipzig.de
**Corresponding author*

Key words: Mars 105 space-flight experiment, urine protein expression, salt balance, molecular processes, bioinformatics analysis, self organizing maps, visualizing high throughput data

Motivation and aim: Long-term space travel simulation experiments enabled to discover new aspects of human metabolism. Detailed proteomics data were collected during the Mars105 isolation experiment potentially enabling a deeper insight into the molecular processes involved. *Methods and algorithms*: Machine learning using self organizing maps (SOM) in combination with different analysis tools was applied to describe the time trajectory of protein expression. The method portrays the protein expression landscapes enabling a personalized and intuitive view on the physiological state of the volunteers.

Results: The dynamics of urine proteomics was described in terms of trajectories reflecting the time evolution of the protein expression landscapes with individual resolution. The abundance of more than one half of the proteins measured clearly changes in the course of the experiment. The trajectory splits roughly into three time ranges, an early (week 1-6), an intermediate (week 7-11) and a late one (week 12-15).The total protein expression level is maximum in the early time region and then it progressively decreases until the end of the experiment. Regulatory modes associated with distinct biological processes were identified using previous knowledge by applying enrichment and pathway flow analysis. Early protein activation modes can be related to immune response and inflammatory processes, activation at intermediate times to developmental and proliferative processes and late activations to stress and responses to chemicals. These protein expression profiles support previous results about alternative mechanisms of salt storage in an osmotically inactive form paralleled by the activation of immune responses in the context of micro-vascularization.

Conclusions: Our study shows that SOM machine learning in combination with analysis methods of class discovery and functional annotation enable the straightforward analysis of complex proteomics data sets generated by means of mass spectrometry.

SOCIAL INTERACTIONS IN COMMUNITIES OF CONFORMISTS AND NON-CONFORMISTS: A CELLULAR AUTOMATON MODEL

Bukharin K.V.[1], Lavryushev Ya.S.[2], Lashin S.A.*[1]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Municipal gymnasium № 3, Novosibirsk, Russia*
e-mail: lashin@bionet.nsc.ru
**Corresponding author*

Key words: social dynamics, opinion dynamics, zealot, social physics, simulation, mathematical model, cellular automaton

Motivation and Aim: In models of social systems, the main objects are people and their face-to-face relations. The aim of mathematical and computer modeling of social phenomena is the investigation of common regularities, events and processes occuring in communities. In particular, mechanisms of disorder-order transition, e.g. formation of common language or culture, reaching consensus etc. In recent times, the question of influence of zealot groups on the rate of consensus reach is of great interest [1]. In the present study, we have constructed and analyzed the model of opinion dynamics in a community of both conformists and nonconformists. We have also analyzed the influence of zealots on such a dynamics.

Methods and Algorithms: We used cellular automaton modeling approach. Each cell was considered to be an agent (one person). A cell could take the following values: A, B (agent gets corresponding opinions), and ZA, ZB (zealot A, zealot B – agent gets corresponding opinion A or B opinion and can not change it further). Rules of state change for a cell defined in the following manner: a cell observes its neighbors (1st, 2nd layers and so on) and changes its state (opinion) onto majority's opinion (community of conformists) or minority's opinion (nonconformists).

Results: Numerical computations have been performed for models of conformists and nonconformists. The grids NxN (N=100, 1000, 5000) have been considered for square and tore topologies. It has been shown that presence of zealots destabilizes rather the community dynamics of nonconformists than of conformists. However, the consensus is reached much faster in the latters.

Conclusion: Our results have shown that the critical influence of radicals on the rate of consensus reach recently reported in [1] is realized in full in communities, which primarily consist of conformists.

Acknowledgements: The work was supported by the RFBR grant 12-07-00671-a, the Stipend of the President of Russian Federation, the Program of the RAS Presidium №28 "Evolution and biosphere origin".

References:

1. Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., & Szymanski, B. K. Social consensus through the influence of committed minorities // Physical Review E, 2011, 84(1), 011130.

UNSTEADY HEMODYNAMIC SIMULATION OF CEREBRAL ANEURYSMS STENTING

Cherevko A.A.[1, 2]*, Chupakhin A.P.[1, 2], Yanchenko A.A.[2], Krivoshapkin A.L.[3], Orlov K.U.[3]
[1]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Meshalkin Novosibirsk Scientific Research Institute of Circulation Pathology, Novosibirsk, Russia*
e-mail: cherevko@mail.ru

Key words: hemodynamic simulation, cerebral aneurysms, stenting

Motivation and Aim: In this work a stent surgery on a brain aneurysm, changing geometry of a blood vessel and the flow in it, is simulated. This problem is interesting in terms of hydrodynamics and has important applications in medicine;

Methods and Algorithms: The geometry of the vessel is reconstructed by tomographic data. To describe the blood flow Navier-Stokes equations for a viscous, incompressible Newtonian fluid are used. Behavior of the vessel wall is described by the linear theory of elasticity. Numerical calculations are carried out in the package ANSYS in Information and Computing Centre of Novosibirsk State University;

Results: A real operation performed by neurosurgeons of Meshalkin Novosibirsk Institute of Circulation Pathology is modeled. Three geometrical configurations of the aneurysm are considered: before surgery, after surgery, and for a control examination (in a year). For all configurations hydrodynamic and mechanical parameters are calculated: velocity, streamlines, WSS, energy loss, the stress and deformations of the wall. A comparative analysis of different simulations is performed;

Conclusion: The results obtained are new, as tracked by real clinical data for different models of aneurysms (steady and unsteady flows; heterogeneous model) and various patient conditions (before surgery, immediately after and a year after). It is interesting for development of pre-surgery modeling.

Availability: not available.

Acknowledgements: RFBR 14-01-00036, SB RAS IP №44, RAS department of Mechanics №2.13.4.

UNSTEADY HEMODYNAMIC SIMULATION OF THE BRAIN'S VASCULAR SYSTEM WITH ANEURYSMS

Cherevko A.A.[1, 2]*, Chupakhin A.P.[1, 2], Khe A.K., Vorontsova E.A., Fedorova Y.A., Krivoshapkin A.L.[3], Seleznev P.A.

[1]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia;*

[2]*Novosibirsk State University, Novosibirsk, Russia;*

[3]*Meshalkin Novosibirsk Scientific Research Institute of Circulation Pathology, Novosibirsk, Russia*

e-mail: cherevko@mail.ru

Key words: hemodynamic simulation, cerebral aneurysms, the brain's vascular system

Motivation and Aim: This work presents a mathematical model and numerical simulation results of the blood flow- structural interaction, which occurs in a saccular aneurysms emerging out of the brain's vascular system, using computational domains made of by medical images reconstructions. Hemodynamics play an important role in the mechanisms that govern the initiation, growth, and possible rupture of intracranial aneurysms. The purpose of this work was to objectively characterize these dynamics, classify them, and connect them to aneurysm rupture. This goal has important applications in medicine.

Methods and Algorithms: the computational domain is generated by image reconstruction techniques out of digital imaging and communication in medicine (DICOM) datasets acquired by computed tomography (CT). Navier-Stokes equations for a viscous, incompressible Newtonian fluid are used for describing the blood flow. Numerical calculations are carried out in the package ANSYS in information and computing centre of Novosibirsk State University;

Results: There was the solving of unsteady hemodynamic simulation of the brain's vascular system with aneurysms. A real operation performed by neurosurgeons of Meshalkin Novosibirsk Institute of Circulation Pathology is modeled. Area of influence of the aneurysms on the blood flow is allocated. As a result of 1-way FSI calculation the stresses in the vessel wall of aneurism and near it were determined.

Conclusion: The mathematical model accounts for the nonlinear behavior of the blood vessels and the one-way coupling between the hemodynamic and structural models. Besides the realistic computation domain – out of medical CT images – pulsating flow conditions and deformable vessel walls are the main features of the model presented in this paper.

Availability: not available.

Acknowledgements: RFBR 14-01-00036, SB RAS IP №44, RAS department of Mechanics № 2.13.4.

CHEMPAK SOFTWARE PACKAGE: NUMERICAL MODELING OF DIRECT AND INVERSE PHARMACOKINETICS PROBLEMS

Chernykh I.G.[1]*, Kabanikhin S.I.[1], Voronov D.A.[2]
[1]Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia;
[2]Novosibirsk State University, Novosibirsk, Russia
e-mail: chernykh@ssd.sscc.ru
**Corresponding author*

Key words: chemical kinetics, pharmacokinetics, high performance computing

Motivation and aim: Modeling of the drug diffusion in the human body is important for the development of new drugs. The currently available dual-chambers pharmacokinetics models do not meet the needs of the rapidly growing pharmaceutical industry. Each multichamber pharmacokinetics problem can be represented by a system of ordinary differential equations. ChemPAK [1] software package was adapted for solving these systems.

Methods and algorithms: Authors propose new solvers for ChemPAK package. Solvers for initial value problem (IVP) for ordinary differential equation systems (ODE) proposed for supercomputers with Intel Xeon Phi and NVIDIA Tesla/Kepler accelerators. These solvers are based on BDF (Backward Differentiation Formula) method, where a nonlinear system must be solved (approximately) at each time step, for which Newton iteration is used, where linear systems must be solved.

Results: ChemPAK was tested on five chambers pharmacokinetics problem, parallel solver was tested on different chemical kinetics systems [3-5]. The results of testing are in a good accordance with experimental data.

Conclusion: There is a new modification of ChemPAK software package presented in this paper. Adapted version of ChemPAK has been tested on five chambers pharmacokinetics problem. The results of parallel solver tests are in a good accordance with experimental data.

References:

1. I. Chernykh, O. Stoyanovskaya, O. Zasypkina. (2009) ChemPAK Software Package as an Environment for Kinetics Scheme Evaluation, Chemical Product and Process Modeling, 4: paper 3, 1-15.
2. I. Chernykh, M. Antonova. (2013) Chemical kinetics scheme evaluation, Proc. Of The 6th Asia-Pacific Congress On Catalysis Taipei Taiwan 2013.
3. V.N. Snytnikov, T.I. Mischenko, Vl.N. Snytnikov, I.G. Chernykh. (2009) A reactor for the study of homogeneous processes using laser radiation energy, Chemical Engineering Journal, 150: 231-236.
4. V.N. Snytnikov, T.I. Mischenko, Vl.N. Snytnikov, I.G. Chernykh. (2012) Physicochemical processes in a flow reactor using laser radiation energy for heating reactants, Chemical Engineering Research and Design, 90, 11: 1918-1922.
5. V.N. Snytnikov, T.I. Mischenko, Vl.N. Snytnikov, I.G. Chernykh. (2013) Autocatalytic dehydrogenation of propane, Res. Chem. Intermed., 39: 1-12.

COMPREHENSIVE STUDY OF HEMODYNAMICS OF CEREBRAL VESSELS IN THE PRESENCE OF PATHOLOGIES

Chupakhin A.P.[1, 2]*, Cherevko A.A.[1, 2], Khe A.K.[1, 2], Chebotnikov A.V.[1], Krivoshapkin A.L.[3], Orlov K.Yu.[3], Panarin V.A.[3]
[1]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Meshalkin Novosibirsk Research Institute of Circulation Pathology, Novosibirsk, Russia*
e-mail: alexander190513@gmail.com
**Corresponding author*

Key words: cerebral hemodynamics, arteriovenous malformation, cerebral aneurysm, mathematical modeling

Motivation and Aim: In this work we present the results of a comprehensive study of hemodynamics of cerebral vessels in the presence of anomalies, such as arteriovenous malformations (AVM) and cerebral aneurysms (CA). The aim of this study is to construct and investigate mathematical models governing blood flow in the cerebral vessels, to construct blood flow parameters, which can determine the occurrence of potentially dangerous complications.

Methods and Algorithms: The researched carried out is complex and includes measurement of blood flow parameters (velocity and pressure) directly during neurosurgical operations at Meshalkin Novosibirsk Research Institute of Circulation Pathology; construction of mathematical models based on the data obtained during examinations; 3D computer simulations, and calculation of parameters important from a medical point of view.

Results: A method of intravascular measurements of blood flow hydrodynamic parameters is developed. The procedure was used during endovascular treatments of AVMs and CAs. A mathematical model governing the flow pressure is constructed. The model parameters are determined from the clinical data. Patient-specific simulations of aneurysm stenting were performed. The results obtained are in good accordance with medical assessments of the operations.

Conclusion: A new technique of intravascular measurement is introduced and implemented. The procedure allows neurosurgeons to assess the hemodynamic parameters of the blood flow in addition to visual estimation of a lesion. The mathematical models developed and patient-specific 3D computer simulations performed in this work reproduce qualitatively correct changes of hemodynamics parameters measured during examination. This can be used for preoperative analysis and treatment strategy decisions.

Acknowledgements: This work was supported by Presidium of the Siberian Branch of the Russian Academy of Sciences (Integration Project 44), the Department of Energetics, Engineering, Mechanics and Control Processes of the Russian Academy of Sciences (Program 2.13.4), Russian Fund for Basic Research (Project 14-01-00036).

HIGH RESOLUTION COMPUTATIONAL MODELS FOR BIOELECTRIC IMPEDANCE ANALYSIS

Danilov A.A.[1]*, Kramarenko V.K.[2], Salamatova V.Yu.[2], Yurova A.S.[3]
[1]*Institute of Numerical Mathematics RAS, Moscow, Russia;*
[2]*Moscow Institute of Physics and Technology, Dolgoprudny, Russia;*
[3]*Lomonosov Moscow State University, Moscow, Russia*
e-mail: a.a.danilov@gmail.com
**Corresponding author*

Key words: segmentation, tetrahedral mesh, finite element method, bioimpedance

Motivation and Aim: Bioelectrical impedance analysis (BIA) is commonly used for body composition and abdominal adiposity assessment in clinical medicine, dietology and sports medicine. The computational analysis of the existing measurement schemes is essential for accurate data interpretation and the development of new efficient electrode schemes.

Methods and Algorithms: We developed a numerical model for computation of the human body bioelectrical impedance for low frequency electric signals [1]. Initial segmentation was performed using ITK-SNAP package from the Visible Human Project data [2]. Unstructured tetrahedral mesh is generated by CGAL library. Ani3D package is used for finite element discretization.

Results: We simulated BIA measurements at the electrical current frequency 50 kHz. Current density fields were calculated from the finite-element model, and the corresponding sensitivity field distributions were obtained for various configurations of electrode pairs [3]. The combined high sensitivity zones cover the whole human body.

Conclusion: In this study we presented the segmentation and mesh generation technology for human body bioelectrical impedance modeling. We created a tetrahedral mesh and obtained the sensitivity field distributions for the conventional tetrapolar measurement scheme as well as for the segmental BIA as represented by the eight- and ten-electrode configurations. Our data provide accurate interpretation of the results of segmental BIA.

Availability: Segmentation package ITK-SNAP is freely available at http://itksnap.org/. Mesh generation library CGAL is freely available at http://cgal.org/. Dicretization package Ani3D is freely available at http://sourceforge.net/projects/ani3d/.

Acknowledgements: This work was partially supported by the RFBR grant 14-01-14-01-00830 and Russian Federation President Grant for Government Support of Young Russian Scientists MK–3675.2013.1.

References:

1. A.A. Danilov et al. (2012) Modelling of bioimpedance measurements: unstructured mesh application to real human anatomy, *Russ. J. Num. Anal. Mat. Mod.*, **27:** 431–440.
2. A.A. Danilov et al. (2013) Mesh generation and computational modeling techniques for bioimpedance measurements: an example using the VHP data, *J. Phys.: Conf. Series*, **407:** 012004.
3. A.A. Danilov et al. (2013) Sensitivity field distributions for segmental bioelectrical impedance analysis based on real human anatomy, *J. Phys.: Conf. Series*, **434:** 012001.

PATIENT SPECIFIC 3D MODELS: SOME GENERATION TECHNIQUES

Danilov A.A.[1], Yurova A.S.[2]*
[1]*Institute of Numerical Mathematics RAS, Moscow, Russia;*
[2]*Lomonosov Moscow State University, Department of Computational Mathematics and Cybernetics, Moscow, Russia*
e-mail: alexandra.yurova@gmail.com
**Corresponding author*

Key words: segmentation, patient-specific models

Nowadays computerized models of real human anatomy are widely used in various research areas such as nuclear medicine, radiography, radiation protection and others. Especially, impedance simulation studies require solving a number of well-defined computational problems using highresolution anatomically accurate 3D models. Our aim is to describe the possible algorithms for creation of patient specific high-resolution 3D models.

Our group has already created a fullsize segmented models of male and female human bodies (see, e.g., 2]) . We propose transformation of the segmented reference model[1]. User selects several control planes, and tries to adapt the reference model in these planes. Model adaptation in the plane is based on piecewise affine mapping defined by the set of control points, which usually represent the anatomical and geometrical features of human body. The size and the form of the individual tissues may be varied using the adaptation.

In some cases it is easier to use methods of semi-automatic segmentation, which are widely represented, e.g. in ITK library. The library contains various modifications of region growing methods, which start from a seed region. Region growing algorithms vary depending on the criteria used to decide whether a voxel should be included in the region or not, the type connectivity used to determine neighbors and the strategy used to visit neighboring voxels. Also library provides a wide choice of image filters, which may be used for preprocessing.

The method applied for segmentation usually depends on the structure of the area to be segmented and the quality of the input medical images (CT or MRI). Sometimes it is convenient to use some combination of segmentation tools. The fully automatic segmentation of human body images still remains a problem.

References:

1. Danilov A.A., Kramarenko V.K., Nikolaev D.V., Yurova A.S. Personalized model adaptation for bioimpedance measurements optimization // Russ. J. Numer. Anal. Math. Modelling. 2013. V. 28, № 5. P.459-470. DOI: 10.1515/rnam-2013-0025.
2. A.A. Danilov, D.V. Nikolaev, S.G. Rudnev, V.Yu. Salamatova and Yu.V. Vassilevski, Modelling of bioimpedance measurements: unstructured mesh application to real human anatomy. Russ. J. Numer. Anal. Math. Modelling (2012) 27, No 5, 431–440.

NUMERICAL SIMULATION OF BLOOD FLOW IN THE VASCULAR NETWORK WITH PATHOLOGIES OR IMPLANTS

Dobroserdova T.K.
Institute of Numerical Mathematics RAS, Moscow, Russia
e-mail: DobroserdovaTK@gmail.com

Key words: model of global blood circulation, 1D-3D model, the state equation, boundary conditions, vascular pathology or implant

Motivation and Aim: Cardiovascular diseases are one of the main causes of human death. 1D model of global blood circulation helps to describe hemodynamics in the healthy vascular network. We develop new methods to take into consideration influence of endovascular implants or pathologies by such model.

Methods and Algorithms: The 1D model of global blood circulation is based on the mass and momentum conservation laws. The third equation of this model, so-called the state equation, represents the dependence between transmural pressure and the area of vessel cross section. We propose two approaches to take into consideration influence of endovascular implants or pathologies by the model of global blood circulation.

The first approach is based on modification of the state equation. The state equation is an empirical function for healthy vessel. Endovascular implants or pathologies change elastic properties of the vessel wall. The state equation for such vessels changes too. We suggest its modification using fiber or fiber-spring model of the elastic vessel wall [1].

The second approach is multiscale: 3D and 1D models of blood flow are coupled. We need to prescribe coupling boundary conditions. Most of standard choices lead to the energy inconsistency in the coupled model. It is not clear if the cumulative energy dissipates in time or not. We offer to use new boundary condition to provide the correct energy balance of the coupled model [2]. It demands the continuity of the linear combination of the fluid flux and the energy flux between 3D boundary Γ_{out} and 1D boundary $x=d$:

$$\bar{p}\int_{\Gamma_{out}} \mathbf{u}\cdot\mathbf{n}\,d\mathbf{s} + \frac{\rho}{2}\int_{\Gamma_{out}} |\mathbf{u}|^2(\mathbf{u}\cdot\mathbf{n})\,d\mathbf{s} = \left(\bar{p}S\bar{u} + \frac{\rho}{2}S\bar{u}^3\right)\Big|_{x=d}$$

where **u** is the velocity of blood in 3D; **n** – outward normal vector; $\bar{u}$, S, $\bar{p}$ are the velocity, the area of cross section and the pressure in 1D; ρ is the density of blood. This condition should be combined with the continuity of normal stress. The correct energy balance is valid for the coupled model with such boundary conditions. The problem can be easily decoupled with splitting methods into the separate 1D problem and the 3D problem with usual inflow-outflow boundary conditions on every time step.

Results: Both methods were used in numerical experiments. We simulated the blood flow in the vascular network with atherosclerotic plaque and in the vena cava with implanted cava-filter. The results of experiments correspond to real data.

Conclusion: Two approaches for blood flow simulation in the vessel network with pathologies or implants have been developed. The results of numerical experiments conform to medical data.

Availability: For 3D calculations we used free software package ANI3D: Advanced Numerical Instruments ({{http://sourceforge.net/projects/ani3d}}).

Acknowledgements: This work has been supported in part by Grants 14-01-00830, 3675.2013.1.

References:

1. Y. Vassilevski, S. Simakov, V. Salamatova, Y. Ivanov, T. Dobroserdova. (2013) Vessel wall models for simulation of atherosclerotic vascular networks. Mathematical Modelling of Natural Phenomena, 6(7): 82–99.
2. T. Dobroserdova, M. Olshanskii. (2013) A finite element solver and energy stable coupling for 3D and 1D fluid models. Computer Methods in Applied Mechanics and Engineering, 259: 166 – 176.

RECONSTRUCTION CYCLIC SEQUENCES FROM THEIR CIRCULAR DISTANCES MULTISET

Fomin E.S.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: fomin@bionet.nsc.ru

Key words: sequencing of cyclic peptides, mass spectra, beltway problem

Motivation and Aim: Some of the most effective antibiotics (i.e., Vancomycin and Daptomycin) are cyclic peptides produced by non-ribosomal biosynthetic pathways. The sequencing of cyclic peptides by using the mass spectrometry technique can be formalized as a problem of reconstruction of point sets from their circular distances multiset [1].

This problem arises in a lot of fields in engineering and applied physics[2,3], and has confounded researchers for over 80 years. It, called as the beltway problem, is one of the few fundamental problems that are neither known to be NP-hard nor solvable by polynomial-time algorithms [3]. Up to now only the $O(n^n \log n)$-time reconstruction algorithms have been presented in literature. These algorithms are capable to reconstruct the sequences which contains up to several tens of elements.

Results: In the study, we give a graph algorithm for the beltway problem that in practice is capable of handling almost all distance multisets of up to 25 thousands points, i.e. the algorithm is capable to restore the cyclic sequences of up to 160 points.

Availability: on the demand.

References:

1. H. Mohimani et al. (2011) Multiplex De Novo Sequencing of Peptide Antibiotics, Journal of computational biology, 18:11 1371–1381.K. Jaganathan, B. Hassibi. (2013) Reconstruction of integers from pairwise distances. Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference, 5974 – 5978.
2. P.Lemke, S.S.Skiena, W.D.Smith. (2003) Reconstructing Sets From Interpoint Distances. Discrete and Computational Geometry. Algorithms and Combinatorics. 25: 597-631.

1D MODELLING OF DIFFERENT TIME REGIMES OF ENHANCED EXTERNAL COUNTERPULSATION

Gamilov T.M.*, Simakov S.S.
Moscow Institute of Physics and Technology, Dolgoprudny, Russia
e-mail: gamilov@crec.mipt.ru
**Corresponding author*

Key words: cardiovascular system, enhanced external counterpulsation, EECP, hemodynamic, vessels network, cardiovascular modelling, elastic tube

Motivation and Aim: Enhanced external counterpulsation (EECP)is a non-invasive procedure that involves surrounding patient's legs with inflatable cuffs that are pressurized and depressurized out-of-phase with systole. It is used for the treatment of patients with different heart diseases [1,]. Despite some success [2], the mechanisms by which EECP improves cardiac function remain unclear. This leads to problems in defining some EECP parametrs for each particular case. The aim of this work is to perform numerical simulations of different EECP regimes and estimate their impact on hemodynamics in coronary region.

Methods and Algorithms: The model viscous incompressible fluid flow through the network of elastic tubes described in [3] was used. This model was validated by a number of methods, including shock wave formation study [4], and enhanced with autoregulation and venous valves models. Heart cycle was considered to be constant and equals 1 s.

Results: Different time regimes of EECP were simulated. Regimes are denoted by the time inside a heart cycle when the pressurization starts: 0.25 s, 0.3 s, 0.35 s, 0.4 s. The increase in average blood pressure in a terminal coronary artery for each regime is 29.2 ± 0.2 %. The emptying effectiveness of leg vessels is different for each regime: 13 %, 16 %, 19 % and 22 % respectively.

Conclusion: Results show that the impact of different EECP regimes on blood circulation in coronary vessels is approximately the same as long as pressurization is out-of-phase with systole. Regimes with a "late start" provided better emptying effectiveness thus , providing higher venous return. We conclude that later start is more effective in terms of assisting in heart work.

Acknowledgements: This work was supported by grant RSCF 14-21-00107.

References:

1. E.T. Ozawa, K.E. Bottom, X. Xiao, R.D. Kamm. (2001) Numerical simulation of enhanced external counterpulsation, *Ann Biomed Eng* **29:** 284-297. W.E. Lawson et al. (1992) Efficacy of enhanced external counterpulsation in the treatment of angina pectoris, *Am J Cardiol* **70**: 859-862.
2. S.S. Simakov, A.S. Kholodov. (2009) Computational study of oxygen concentration in human blood under low frequency disturbances, *Math Mod Comp Sim* **1(2)**: 283-295.
3. S. Canic, E.H. Kim. (2003) Mathematical analysis of the quasilinear effects in a hyperbolic model blood flow through compliant axi-symmetric vessels, *Math Meth Appl Sci* **26**: 1161-1186.

SOLUTION OF THE INVERSE LIGHT-SCATTERING PROBLEM FOR CHARACTERIZATION OF RED BLOOD CELLS

Gilev K.V.[1, 2]*, Yurkin M.A.[1, 2], Chernyshev A.V.[1, 2], Maltsev V.P.[1, 2]
[1]*Voevodsky Institute of Chemical Kinetics and Combustion SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia*
e-mail: kgilev@gmail.com
**Corresponding author*

Key words: inverse problem, light scattering, red blood cells

Motivation and Aim: Hematological analyzers, that are widely used in clinical diagnostics, measure volume of red blood cells and hemoglobin concentration. Determination of red blood cells shapes and index of sphericity can be used in detailed analysis of human blood for diagnosis of various diseases.

Methods and Algorithms: Scanning Flow Cytometer (SFC) is a novel device measuring light scattering patterns (LSP) of single particles in flow [1]:

$$I(\theta) = \frac{1}{2\pi}\int_0^{2\pi} \left(S_{11}(\theta,\varphi) + S_{14}(\theta,\varphi)\right)\mathrm{d}\varphi\,,$$

where S – the Mueller matrix, θ and φ – polar and azimuthal scattering angles, $\theta \in [10°, 70°]$. Using this approach blood sample can be analyzed apiece with independent analysis of every particle.

We consider an inverse problem consisting in determination of shape parameters of red blood cells from measured LSP. We model a red blood cell as biconcave disc described by 5 parameters (volume V, surface area S, spontaneous curvature C_0, orientation angle β, refractive index n). The particular shape of red blood cell is determined from the minimization of deformation energy with fixed V, S, and C_0 [2].

We used Discrete Diploe Approximation (DDA) method for solution of the direct light-scattering problem [3].

Results: In the experiment we measure 10^2–10^3 cells per second. Due to the fact that solution of direct light-scattering problem takes about 1 minute at personal workstation, direct fitting is not feasible. Instead the inverse solution is realized by nearest-neighbor interpolation on the preliminary calculated database with $2\cdot10^6$ records.

References:

1. V.P. Maltsev, "Scanning flow cytometry for individual particle analysis," *Rev. Sci. Instrum.,* **71**: 243–255 (2000).
2. H.J. Deuling et W. Helfrich. The curvature elasticity of fluid membranes : A catalogue of vesicle shapes, J. Phys. France, 37: 1335-1345 (1976).
3. Yurkin M.A. and Hoekstra A.G. The discrete-dipole-approximation code ADDA: capabilities and known limitations, JQSRT, 112: 2234–2247.

COMPUTATION OF DRUG INTERACTIONS AND SIDE EFFECTS

Hofestädt R.
Tobias Hoppe and Alban Shoshi Bielefeld University
e-mail: hofestae@techfak.uni-bielefeld.de

For most of the medical drugs which are in practical use, the molecular effect and its metabolic influence is still an open question. Today, methods of Integrative Bioinformatics can help understand the molecular processes and biochemical mechanisms of drugs. One result of such studies is the reducing of side effects. The development of relevant pharmaceutical information systems supporting this kind of work is extremely important because in countries like Germany, each year the deaths of around 30000 patients are attributed to drugs. Overall adverse drug reactions is one of the most common causes of death in industrialized countries. Nowadays, useful medical, biochemical and patient data is available to avoid such effects. Therefore, empirical data based on clinical studies for the approval and monitoring of drugs and molecular information systems is available. Based on these information systems methods of Integrative Bioinformatics allow the fusion of pharmaceutically relevant information. To solve this task, we can use medical information systems like the RED LIST or ABDAMED, which present information to the medical and pharmaceutical personal about the effects and side effects of any drug which is on the market. The data of such information systems represent well-collected empirical data and information about the agents of any drug. Agents represent the bridge from the drug into the molecular world. That means that by having information about human pathways, human genes and human omic data we can try to identify the relevant human networks where these agents react. At this point, we can fuse the molecular data coming from different databases like KEGG etc. Furthermore, special databases representing the molecular drug effect in relation to side effects are available. Overall, currently we are able to identify drug side effects based on the present medical and molecular niveau.

This talk presents the idea of our web-based decision support system GraphSAW, which analyzes and evaluates drug interactions and side effects based on data from two commercial and two freely available molecular databases. The system is able to analyze single and combined drug-drug interactions, drug-protein interactions as well as single and cumulative side effects. In addition, it is possible to map diseases and drugs on metabolic pathways. GraphSAW provides a comprehensive and improved medication analysis which reduces the risk of medication errors and a number of drastic side effects.

An application case of 20 leading drugs is presented in order to demonstrate the functionality of the system under real conditions.

References:

1. ABDAMED [Internet database]. Eschborn/Taunus, Vertriebsgesellschaft Deutscher Apotheker mbH. c2013 [cited 2013 Jan 01]. Available at: http://www.abdata.de
2. R. Kaushal, K. Shojania, D. Bates. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. Arch Intern Med, 163(12): 1409-1416, 2003
3. T. Töpel, B. Kormeier, A. Klassen, R. Hofestädt. BioDWH: a data warehouse kit for life science data integration. J Integr Bioinform, 5(2), 2008

THE IDENTIFICATION AND REFINEMENT OF PARAMETERS OF MATHEMATICAL MODELS IN IMMUNOLOGY

Ilyin A.I.[1], Kabanikhin S.I.[2, 3], Krivorotko O.I.[2, 3]*
[1]*The Republican State Enterprise «Scientific Center for Anti-Infectious Drugs», Almaty, Kazakhstan;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Computational Mathematics and Mahemational Geophysics SB RAS, Novosibirsk, Russia*
e-mail: krivorotko.olya@mail.ru
**Corresponding author*

Key words: immunology, inverse problems, numerical methods, optimization methods

Motivation and Aim: Major part of mathematical models in biology is described by systems of nonlinear differentional equations. Currently the mathematical modeling of immunological systems based on numerical solution of ordinary differentional equations (ODEs) grows rapidly. Developed immunological models characterizes by its own intrinsic parameters. Such coefficients describe pathogen, features of immune response, *etc*. That's why it is necessary to identify these parameters for getting information about nature of disease, immune response, susceptibility to specific medication, *etc*.

Methods and Algorithms: Inverse problem is solved by different algorithms: Landweber iterations method, conjugate gradient method and Singular Value Decomposition.

Results: We have developed an algorithm for refinement the parameters of immunology models by the example of acute pneumonia. We validate the inverse problem solution with real parameters of disease.

Conclusion: It is shown that physical properties of initial approximations strongly affect on obtained solutions. The results of numerical experiments are presented.

Acknowledgements: This work was supported by Ministry of Education and Science of the Russian Federation and interdisciplinary project of SB RAS 122.

References:

1. A.I. Ilyin, S.I. Kabanikhin, O.I. Krivorotko. The refinement of parameters of immunology models described by nonlinear differential systems. Siberian Electronic Mathematical Reports, 2014 (in Russian).
2. S. Marino*, D.E. Kirschner. The human immune response to Mycobacterium tuberculosis in lung and lymph node. Journal of Theoretical Biology. **227** (2004) 463–486.
3. G.I. Marchuk. Mathematical Modelling of Immune Response in Infectious Diseases. Springer-Science+Business Media, 1997.

INVERSE PROBLEMS FOR DIFFERENTIAL EQUATIONS OF PHARMACOKINETICS AND IMMUNOLOGY

Ilyin A.I.[1], Kabanikhin S.I.[2, 3]*, Krivorotko O.I.[2, 3], Voronov D.A.[2, 3]
[1]*The Republican State Enterprise «Scientific Center for Anti-Infectious Drugs», Almaty, Kazakhstan;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Computational Mathematics and Mahemational Geophysics SB RAS, Novosibirsk, Russia*
e-mail: voronov-dima@mail.ru
**Corresponding author*

Key words: Dynamic differential equations, Pharmacokinetics, Immunolgy, Inverse problems, Numerical methods, Identifiability

Motivation and Aim: Dynamic ordinary differential equations describes variety of processes occurred in day-to-day life. We are turning our attention to two of them. First of these two is the problem of modeling pharmacokinetic processes in human body. Pharmacokinetics deals with kinetics of absorption, distribution, metabolism and excretion of drugs and their corresponding pharmacologic, therapeutic or toxic responses in man and animals. What happens to the drug in the body can be visualized by considering it as being made up of a large number of compartments, each of which has a volume where the drug is well mixed. Drug is then transferred between these compartments, either transported by the blood from one to another, or by passing an interior membrane in some body organ. Second problem connected to mathematical modeling of immunological systems. Developed immunological models characterizes by its own intrinsic parameters. Such coefficients describe pathogen, features of immune response, etc. That's why it is necessary to identify these parameters for getting information about nature of disease, immune response, susceptibility to specific medication, etc.

Methods and Algorithms: In both cases we have dynamic system described by a system of ordinary differential equations. Arising direct problems are solved using Runge-Kutta method while inverse problems are solved by different numerical methods such as Landweber iterations method, Newton-Kantorovich method and Singular Value Decomposition.

Results:. An algorithm for solving inverse problem in case of n-compartment is developed. The results of numerical experiments are presented. Also by the axample of acute pneumonia an algorithm for refinement parameters of immunolgy models.

Conclusion: It is shown that physical properties of initial approximations strongly affect on obtained solutions.

Acknowledgements: This work was supported by Ministry of Education and Science of the Russian Federation and interdisciplinary project of SB RAS 122.

References:

1. Ilyin A., Kabanikhin S.I. , Nurseitov D. ., Nurseitova A.T., Asmanova N.A., Voronov D.A., Bakytov D. Analysis of ill-posedness and numerical methods of solving a nonlinear inverse problem in pharmacokinetics for the two-compartmental model with extravascular drug administration. Journal of Inverse and Ill-Posed Problems. Volume 20, Issue 1, Pages 39–64, ISSN (Online) 1569-3945, ISSN (Print)0928 -0219, DOI: 10.1515/jip-2012-0202, March 2012.
2. Simeone Marino, Denise E. Kirschner. The human immune response to Mycobacterium tuberculosis in lung and lymph node. Journal of Theoretical Biology 227 (2004) 463–486.

NUMERICAL SOLUTIONS OF INVERSE PROBLEM OF PHARMACOKINETICS. IDENTIFIABILITY OF COMPARTMENTAL MODELS

Ilyin A.I.[1], Kabanikhin S.I.[2, 3], Voronov D.A.[2, 3]*
[1]*The Republican State Enterprise «Scientific Center for Anti-Infectious Drugs», Almaty, Kazakhstan;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Computational Mathematics and Mahemational Geophysics SB RAS, Novosibirsk, Russia*
e-mail: voronov-dima@mail.ru
**Corresponding author*

Key words: Pharmacokinetics, Inverse problems, Numerical methods, Identifiability

Motivation and Aim: Pharmacokinetics deals with kinetics of absorption, distribution, metabolism and excretion of drugs and their corresponding pharmacologic, therapeutic or toxic responses in man and animals. What happens to the drug in the body can be visualized by considering the body as being made up of a large number of compartments, each of which has a volume where the drug is well mixed. Drug is then transferred between these compartments, either transported by the blood from one to another, or by passing an interior membrane in some body organ. We can visualize this whole process as a dynamic system described by a system of ordinary differential equations.

Methods and Algorithms: We can visualize this whole process as a dynamic system described by a system of ordinary differential equations. Parameter identifiability analysis for dynamic system ODE models addresses the question of which unknown parameters can be quantifies from given input-output data. The linear compartment models that we focus on in this report are never identifiable, except in the trivial case of a model with only one compartment. This forces us to look for identifiable reparametrization of our model. Inverse problem is solved by different algorithms: Landweber iterations method, Newton-Kantorovich method and Singular Value Decomposition.

Results: In this report we consider scaling reparametrization that is obtained by replacing an unobserved variable by a scaling version of itself, and updating coefficients accordingly. Also an algorithm for solving inverse problem in case of n-compartment is covered in this report.

Conclusion: The question of choosing initial approximations is covered in this report. It is shown that physical properties of initial approximations strongly affect on obtained solutions. The results of numerical experiments are presented.

Acknowledgements: This work was supported by Ministry of Education and Science of the Russian Federation and interdisciplinary project of SB RAS 122.

References:

1. Ilyin A., Kabanikhin S.I., Nurseitov D.B., Nurseitova A.T., Asmanova N.A., Voronov D.A., Bakytov D. Analysis of ill-posedness and numerical methods of solving a nonlinear inverse problem in pharmacokinetics for the two-compartmental model with extravascular drug administration. Journal of Inverse and Ill-Posed Problems. Volume 20, Issue 1, Pages 39–64, ISSN (Online) 1569-3945, ISSN (Print)0928 -0219, DOI: 10.1515/jip-2012-0202, March 2012.
2. John G. Wagner. Linear Pharmacokinetic Equations Allowing Direct Calculation of Many Needed Pharmacokinetic Parameters from the Coefficients and Exponents of Polyexponential Equations Which Have Been Fitted to the Data, Journal of Pharmacokinetics and Biopharmaceutics, Vol. 4, No. 5, 1976.

THE IDENTIFICATION AND REFINEMENT OF PARAMETERS OF MATHEMATICAL MODELS IN IMMUNOLOGY

Ilyin A.I.[1], Kabanikhin S.I.[2, 3], Krivorotko O.I.[2, 3]*
[1]*The Republican State Enterprise «Scientific Center for Anti-Infectious Drugs», Almaty, Kazakhstan;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Computational Mathematics and Mahemational Geophysics SB RAS, Novosibirsk, Russia*
e-mail: krivorotko.olya@mail.ru
**Corresponding author*

Key words: immunology, inverse problems, numerical methods, optimization methods

Motivation and Aim: Major part of mathematical models in biology is described by systems of nonlinear differentional equations. Currently the mathematical modeling of immunological systems based on numerical solution of ordinary differentional equations (ODEs) grows rapidly. Developed immunological models characterizes by its own intrinsic parameters. Such coefficients describe pathogen, features of immune response, *etc*. That's why it is necessary to identify these parameters for getting information about nature of disease, immune response, susceptibility to specific medication, *etc*.

Methods and Algorithms: Inverse problem is solved by different algorithms: Landweber iterations method, conjugate gradient method and Singular Value Decomposition.

Results: We have developed an algorithm for refinement the parameters of immunology models by the example of acute pneumonia. We validate the inverse problem solution with real parameters of disease.

Conclusion: It is shown that physical properties of initial approximations strongly affect on obtained solutions. The results of numerical experiments are presented.

Acknowledgements: This work was supported by Ministry of Education and Science of the Russian Federation and interdisciplinary project of SB RAS 122.

References:

1. A.I. Ilyin, S.I. Kabanikhin, O.I. Krivorotko. The refinement of parameters of immunology models described by nonlinear differential systems. Siberian Electronic Mathematical Reports, 2014 (in Russian).
2. S. Marino*, D.E. Kirschner. The human immune response to Mycobacterium tuberculosis in lung and lymph node. Journal of Theoretical Biology. **227** (2004) 463–486.
3. G.I. Marchuk. Mathematical Modelling of Immune Response in Infectious Diseases. Springer-Science+Business Media, 1997.

ANDSYSTEM: ASSOCIATIVE NETWORK DISCOVERY SYSTEM FOR AUTOMATED LITERATURE MINING IN THE AREA OF BIOLOGY

Ivanisenko V.A.*, Saik O.V., Tiys E.S., Ivanisenko T.V., Demenkov P.S.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;
PBSoft LLC, Novosibirsk, Russia
e-mail: salix@bionet.nsc.ru
**Corresponding author*

Key words: text mining, data mining, ANDSystem, ANDVisio, ANDCell, gene networks, pathways

Motivation and Aim: The number of publications in the areas of biology, medicine, and biotechnology increases with an alarming rate that makes imperative computer-based analysis. To date, over 20 million abstracts highly relevant to biology and medicine are stored in the PubMed database and the number keeps increasing. To address the confounding problem of extraction of information on molecular-genetic objects from texts worldwide, approaches based on algorithms, from the simplest, such as search for co-occurrence of the biological object names in texts, to complex integrating linguistic-semantic analysis, and machine learning have been suggested. Previously we have developed the ANDSystem package that incorporates utilities for automated extraction of knowledge from Pubmed published scientific texts and databases (Demenkov et al, 2011). However, currently existing text mining systems can not completely extract all the information contained in scientific publications. At the same time, it can be assumed that the combined use of different automated text analysis systems will reduce the loss of information. Another significant disadvantage of the existing systems is too generalized description of interactions between biological objects at which they lose their meaning. In most systems, a number of different interactions types is limited by 2-3 only, including the co-occurrence of the names of objects in the publications and protein-protein interaction. At the same time, along with the problem of lack of information there is another side of the problem arising in the reconstruction of any particular molecular genetic network using automated methods as well: the "information explosion". It is well known that in the case when the reconstruction process of the network includes the addition of all genes for which there is information that they interact with the started set of genes, very often the "galaxy" can appear on the second or third iteration. Thus, the goal of this work was a development of a new version of ANDSystem providing a more complete extraction of knowledge, on the one hand, and automated check of the specificity of networks features to the studied phenotypic trait or biological process limiting uncontrolled growth of the network during its reconstruction, on the other hand.

Results: ANDSystem was developed for the purpose of scanning literature for extracting relationships between diseases, pathways, proteins, genes, microRNAs and metabolites. The ANDSystem incorporates utilities for automated extraction of knowledge from Pubmed published scientific texts and analysis of factographic databases. The ANDCell database contains information on molecular-genetic events retrieved from texts and databases. ANDVisio is a new user's interface to the ANDCell database stored on the remote server. ANDVisio provides graphic visualization, editing and search features as well as possibilities to save an associative gene networks in different formats resulting from user's request. The associative gene networks describe semantic relationships between molecular-genetic objects (proteins, genes, metabolites and others), biological processes, and diseases.

ANDVisio is provided with various tools supporting automated reconstruction of associative networks with taking into account the specific relation of objects in the network to the studied phenotypic trait or biological process limiting uncontrolled network expansion. The ANDSystem can assist in the interpretation of complex multifactorial experimental data.

Work was supported in part by a project EU FP7 № 260429.

References:

1. P.S.Demenkov et al. (2011) ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem, In silico biology, 11:149-61.

PATIENT SPECIFIC RECONSTRUCTION OF VASCULAR NETWORK FOR HEMODYNAMIC MODELING

Ivanov Y.A.*, Pryamonosov R.
Institute of Numerical Mathematics RAS, Moscow, Russia
e-mail: yura-vtm@yandex.ru
**Corresponding author*

Key words: hemodynamics, patient specific, 1D network reconstruction algorithm

Motivation and Aim: A study of vascular diseases impact on hemodynamics may be *based* on mathematical models and numerical simulation. It is supposed to use closed blood circulation model [1]. The 1D blood flow simulation requires 1D core graph representation of the 3D patient specific vascular network. Given a 3D vascular domain extracted from MRI data, one needs an algorithm producing the vascular graph.

Methods and Algorithms: The process could be divided into the following steps: image preprocessing (noise filtering); vessels segmentation; extraction of vascular structure (centerline representation of vessels); topological analysis (graph building). Commercial software Amira [2] has wide functionality and it helps to provide each step of the process, especially in trivial cases. But some vascular domains require more sophisticated approach and it is necessary to have a possibility to improve the segmentation methods. We use open source library VMTK [3]. In particular, was proposed an algorithm providing segmentation of coronary arteries from middle mediastinum CT images. The vascular graph is produced from the set of centerlines by a new algorithm presented in detail in [4].

Results: We perform automatic segmentation for the large legs arteries with basic VMTK methods and for coronary arteries with new algorithm, then produce vascular graph for all the geometries with topological analysis algorithm.

Conclusion: For patient-specific vascular network reconstruction we adopt the open source library VMTK to produce vascular centerlines on the basis MRI/CT data followed by the topological analysis algorithm. The produced vascular graph possesses all necessary geometric data for hemodynamic simulations. We demonstrated applicability of our approach to predictive personalized postsurgical blood flow simulations.

Availability: The information about software mentioned above is available in [2, 3]. All the developed algorithms and methods are academic and not available yet.

Acknowledgements: This work was partially supported by the RFBR grants 11-01-00971, 11-01-00855, Russian Federation President Grant for Government Support of Young Russian Scientists

References:

1. S.S. Simakov, A.S. Kholodov. Computational study of oxygen concentration in human blood under low frequency disturbances. Mat. Mod. Comp. Sim., 1 (2009), 3–295.
2. http://www.amira.com, http://www.vsg3d.com/amira
3. http://www.vmtk.org
4. S. Simakov, T. Gamilov, Yu. Vassilevskii, Yu. Ivanov, P. Kopylov. Patient specific haemodynamics modeling after occlusion treatment in leg. Submitted to Math. Model. Nat. Phenom.

DYNAMIC ASSOCIATION MAPPING BASED ON RANDOM WALK MODEL USING SIMULATED QTLMAS DATA SET

Karacaören B.
Department of Animal Science, Faculty of Agriculture, Akdeniz University, Antalya, Turkey
e-mail: burakkaracaoren@akdeniz.edu.tr

Key words: Association mapping, Random Walk, Kalman Filter, Gibbs Sampling

Motivation and Aim: Mixed model approach with pedigree information commonly employed to detect and correct for genetic relationship in cross sectional genomics research. However longitudinal approach would be more realistic for detecting complex web of interactions of genes for investigation on quantitative traits (Wu and Li, 2006). In this study we assumed that genomic signal over time could be traced by a random walk model.

Methods and Algorithms: We used random walk model and it is given below (Karacaören, 2006).

$$y_t = \alpha_t + \varepsilon_t,\ \varepsilon_t \propto N(0,\sigma_e^2);\ \alpha_{t+1} = \alpha_t + \eta_t,\ \eta_t \propto N(0, A\sigma_n^2) \quad (1)$$

In (1) the first equation is called the observation equation and the second equation is called the state equation. We assumed that observations, y_t, depends on unobservable quantity, α_t, and our aim was to do statistical inference on α_t (states) conditional on genetic relationship matrix, A,. We used 10000 gibbs sampling to predict genetic, permanent environment and residual variance components. For genome wide association analyses; residuals from (1) used as response variable in a single marker regression model for each time point. Data simulated for six Quantitative Trait Locus (QTL) using the pedigree included 2025 individuals with 453 single nucleotide polymorphisms (SNPs) by five time points (Coster et al, 2010).

Results: Residual, genetic and permanent environment variances predicted as 34.46, 24.55, and 26.52 respectively. Correlations between observations and predictions were found to be about 0.85. Number of SNPs based on 1% false discovery rate for time points was; 0, 27, 4, 11, 45. Since random walk model needs more time points to have stable predictions it is expectable that very first time points subject to false positives. Last time point included many signals from true QTLs but false positives as well.

Conclusion: Due to small number of time points our model did not detect all true genomic signals. But still the model may be useful in practice due to its recursive structure. When the longitudinal observations available (daily or monthly for example) the model could predict the on-line genomic signals sequentially without inverting the data matrices.

References:

1. A. Coster et al. (2010) QTLMAS 2009: Simulated dataset. *BMC Proceedings*, **4**(Suppl 1)**:**S3.
2. B. Karacaören (2006) Phenotypic and Genetic Analysis of Functional Traits in Dairy Cattle under Experimental Management System. Dr. Sc. thesis. Swiss Federal Institute of Technology, ETH Zurich, Switzerland.
3. L.R. Wu, M. Lin (2006) Functional mapping – A new tool to study the genetic architecture of dynamic complex traits. Nature Review Genetics **7**: 229-237.

GENE NETWORKS MODELING: SPECIFICATION LANGUAGE

Kazantsev F.V.[1]*, Akberdin I.R.[1], Podkolodnyy N.L.[1, 2, 3], Likhoshvai V.A.[1, 2]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Computational Mathematics and Mathematical Geophysics, SB RAS, Novosibirsk, Russia*
**Corresponding author: kazfdr@bionet.nsc.ru*

Key words: mathematical modeling, gene networks, model specification language, SiBML

Motivation and aim: One of the topic challenge in the Systems Biology is the studying of live system functioning laws. Methods of mathematical modeling and computer simulations are widely used for this task. For the relatively simple models, the universal programming languages or well-known mathematical modeling tools are enough, but for the live system models with the complex hierarchical structure and many interactions in their the development of new technics is required. Growing scale of live system models results in the development of special approaches for description of biological processes functioning that suit the following requirements: applicable for different levels of organization (the genetic, molecular, subcellular and higher levels); the unified description on each level.

Results: We have developed the gene network mathematical model specification language called SiBML. The language and it tools have been developed taking into account a hierarchical conception of biological system models reconstruction. In the framework of the approach the model reconstruction process does in the terms of Elementary Subsystems (ES) which are molecular process or the set of processes that can be modeled as isolated process from the others (for example, molecule transport subsystems, enzymatic synthesis, DNA or RNA synthesis processes etc.). ES in the terms of SiBML can be in two states: full specified (the model where molecules and the law of their interactions are set) and half-specified (the model where only the interactions law for undefined participants are set). The half-specified is the functional template. The complex model reconstruction process on genetic, molecular, subcellular, cellular and others levels of the studied system's organization is performed by independent steps of molecule association with functional templates and their combination with existed one models from the libraries of fully-specified ES. To model the system with complex organization (for example, the metabolic pathways, organelles, whole organism) SiBML has the special scripting functionality. The script specifies hierarchical model-building procedures on basis of existed one ES. This feature provides functionality of local modifications in the model; ES remove/add functionality; generation of model variants. It can be useful for model investigation in the wild conditions, with mutations and/or environment impacts. It helps to keep the time and decreases potential of arising mistakes during the manual model reconstruction process. The analysis of developed mathematical model can be performed using SiBML tool or in the well-known tools as Matlab, Scilab or Octave. The SiBML tool has been also implemented on the HPC cluster of SSCC SB RAS.

Acknowledgements: The work was partially supported by the RFBR (project 13-01-00344), budget projects IC&G SB RAS (VI.61.1.2).

ON SOME ANALYSIS, RECOGNITION AND CLASSIFICATION PROBLEMS OF BIOMETRICAL SEQUENCES IN A CONNECTION WITH COMBINATORIAL OPTIMIZATION PROBLEMS

Kel'manov A.V.
Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia
e-mail: kelm@math.nsc.ru

Key words: biometrical sequences, biomedical signals, analysis, recognition, classification, combinatorial optimization, NP-hard problems, polynomial algorithms with performance guarantee

Motivation and Aim: The research subject of the report includes discrete optimization problems induced by actual problems of biometrical sequences analysis, recognition and classification. The problems are closely connected with processing of such sequences which include repeating or alternating of the information-bearing fragments in the one-dimensional case or information-bearing tuples in the multi-dimensional case. We also study some new optimization problems that are typical for analysis and recognition of sequences which include some combinations on these information-bearing elements. In addition, the problems considered in the report are important and take place in computational geometry, in approximation theory, in mathematical programming and in statistics. The purpose of this report is to review new results on the studying of computational complexity of these problems, and on the substantiating polynomial algorithms with performance guarantee for solution to these problems.

Methods and Algorithms: We present only original mathematical methods and algorithms for solution to the problems.

Results: The main results are following. Some new (previously not studied) and known (weakly studied) problems, and also generalizations and special cases of classical partitioning problems, Euclidean vectors subsets and subsequences search are investigated. NP-hard and polynomial solvable cases of discrete optimization problems and corresponding NP-hard and polynomial solvable problems are found for sequences analysis, their recognition and classification. Exact polynomial and pseudo-polynomial algorithms and effective approximation algorithms with guaranteed accuracy bounds for these problems are constructed.

Conclusion: Here we present new methods and algorithms that form a basis for new effective technologies for processing biometrical sequences and biomedical signals. The algorithms were implemented as software and included into the special computer system QPSLab (Quasi-Periodic Sequences Laboratory). This system is oriented to processing of the data presented as the one- or multi-dimensional number sequences (signals) that contain repeating or alternating informationally valuable patterns and their combinations determining the unique structure of the sequence. Currently the system contains original technologies for solving about fifty typical problems of off-line sequences analysis, their recognition and classification including the problems of detection, restoration, estimation, partition, segmentation, clusterization, identification, decision making, etc.

Acknowledgements: The author was partially supported by Russian Foundation for Basic Research (projects No. 12-01-00090, No. 13–07–00070) and the Target Program of the SB RAS (Integration project No. 21A).

BIOUML: PLUGIN FOR POPULATION-BASED MODELING

Kiselev I.N.[1, 2]*, Kolpakov F.A.[1, 2]
[1]*Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia;*
[2]*Institute of Systems Biology, Ltd, Novosibirsk, Russia*
e-mail: axec@systemsbiology.ru
**Corresponding author*

Key words: non-linear mixed effects model, parameter estimation, R, PharmML

Motivation and Aim: Non-linear mixed effects (NLME) model is an efficient tool for analyzing of population data and estimating of model parameters. It is widely used in pharmacological modeling, but may be applied to a variety of fields. NLME model contains core structural model which is the mathematical model of studying process (e.g. drug dynamics in organism) and suggested distribution of model parameters across population.

Methods and Algorithms: BioUML is an open source integrated Java platform for modeling of biological systems. It provides tools for visual creation and numerical simulation of models with different mathematical formalisms including ODE, PDE, composite models, agent-models etc. "nlme" [1] is a library written in R language implementing log-likelihood approach for NLME models creation and analysis. Core structural model should be supplied as R function which makes usage of sophisticated mathematical models complicated. Library "nlmeODE" [2] allows usage of ODEs inside "nlme". However it has certain restrictions on possible models and textual format complicates using of complex ODE systems.

Results: We have implemented plugin for BioUML platform for population-based modeling. Main features of the plugin are:

1. Graphical notation for structural and population models, facilitating work with detailed and large-scaled models. Both structural model and data about parameters and residual error distribution are represented visually as BioUML diagrams;

2. Parameter estimation is performed by R "nlme" function, supplied (using "rJava") with Java simulator from BioUML. It grants possibility to use any model created in BioUML (including ODE, PDE, algebraic, composite, agent-based) as a structural model. Executing R script is automatically generated on the basis of the diagram;

3. Import and export into PharmML format - rapidly developing standard for pharamacometric models description and exchange between different tools [3].

Conclusions: Developed software combines visual representation of mixed-effects models, modeling and simulation tools in BioUML for structural model simulation, and R "nlme" algorithms for parameter estimation and analyzing of mixed-effects models.

Availability: Plugin is freely available as a part of BioUML software at www.biouml.org.

References:

1. Pinheiro J.C., Bates D.M. Mixed-Effect Models in S and S-PLUS, Springer-Verlag, New York, NY, 2000.
2. Tornoe C.W., et al. Non-linear mixed-effects pharmacokinetic/pharmacodynamic modelling in NLME using differential equations. Computer Methods and Programs in Biomedicien, 2004. 76(1) 31-40.
3. Moodie, S.L., et al. PharmML: The Pharmocometrics Markup Language. Language Specification. 2013. (www.pharmml.org).

COMPLEX DYNAMICS IN SYSTEMS OF ALTERNATIVE mRNA SPLICING: A MATHEMATICAL MODEL

Kogai V.V.[1, 3]*, Fadeev S.I.[1, 3], Khlebodarova T.M.[2], Likhoshvai V.A.[2, 3]
[1]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;*
[2]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[3]*Novosibirsk State University, Novosibirsk, Russia*
e-mail: kogai@math.nsc.ru
**Corresponding author*

Key words: Alternative splicing, mathematical modeling

Motivation and aim: Alternative splicing as a mechanism for the regulation of gene activity is widespread among eukaryotes. Proteins that are formed as a result of this phenomenon, play an important role in all stages of the functioning of living organisms. About 95% of human genes are subjected to alternative splicing. Analysis of the dynamic properties of such systems is the current actual problem of systems biology.

Results: We study a simple mathematical model of alternative splicing, in which two transcription factor isoforms are controlling expression efficiency of own gene through feedback mechanism. Model represented by a system of two differential equations with two delayed arguments. Numerical study of the system depending on the parameters indicates the various modes of operation: quasi-periodic and periodic solutions with period doubling; solutions, unstable on the initial data.

Availability: Further study of the mathematical model will involve consideration of various mechanisms of autoregulation.

Acknowledgements: The work was supported by the RFBR (project 13-01-00344), by the Program of Siberian Branch of the Russian Academy of Sciences (Integration project № 80), budget projects IC&G SB RAS(VI.61.1.2).

APPLICATION OF A BIEXPONENTIAL FORM OF TWO-COMPARTMENT PHARMACOKINETIC MODELS WITH FIRST-ORDER ABSORPTION IN PHARMACOKINETIC ANALYSIS

Koloskov G.*, Asmanova N., Ilyin A.I.
Scientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan
e-mail: gk-v@yandex.ru
**Corresponding author*

Key words: biexponential equation, vanishing exponential, coupled solutions, flip-flop

Motivation and Aim: The purpose of this study is to demonstrate that in compartmental pharmacokinetic (PK) analysis the sole fact, that the observed behavior of the PK curve is found to be biexponential (with exponents λ_1, λ_2) does not provide enough evidence to conclude that the nature of underlying PK model with first-order absorption is one-compartmental with the rate constants of elimination $k_{10} = \lambda_1$ and absorption $k_a = \lambda_2$ (or $k_{10} = \lambda_2$, $k_a = \lambda_1$ in a flip-flop case [1]).

Methods and Algorithms: Comparative analysis of solutions for one- and two-compartment PK models with first-order absorption suggests that exactly the same PK curve would be observed in a special biexponential case of two-compartment model with first-order absorption with vanishing exponent λ_3, when k_{21} approaches λ_3. In this case λ_2 and λ_1 will be interpreted as α and β of the corresponding two-compartment model, and the rate constant of absorption $k_a = \lambda_3$: $\lambda_1 < k_a < \lambda_2$ will satisfy the equation for the one-compartment model, resulting in the same PK curve, while distribution of the amount of the drug would not be uniform, as presumed in case of one-compartment model. This biexponential form of two-compartment model smoothly transforms into 'pure' one-compartment model when k_a approaches λ_2 (or λ_1 in case of flip-flop).

Results: The study suggests an approach for estimation of the maximal potential accumulation of the drug in the peripheral compartment compared to the amount of drug in the main compartment.

Conclusion: Unjustified ignoring of the two-compartmental possibility in case of observed biexponential curve can lead to misinterpretation of the system behavior as uniform, and underestimation of the amount of the drug accumulated in tissue, as well as a misestimation of the PK parameters of the underlying model.

References:

1. Asmanova N., Koloskov G., Ilin A. (2013) Coupled solutions of one- and two-compartment pharmacokinetic models with first-order absorption. J. Pharmacokinet. Pharmacodyn., 40: 229–241, DOI:10.1007/s10928-013-9312-6

INFERRING HYPOTHESES ON PROTEIN-PROTEIN INTERACTIONS DURING EXPERIMENTAL TREATMENT USING PROFILING OF PROTEINS AND PPI DATA FROM PUBLIC DOMAIN

Kolosov M.S.*, Brovchenko P.P.
Southern Federal University, Rostov-on-Don, Russia
e-mail: msk@sfedu.ru
**Corresponding author*

Key words: interactomics, Gene Ontology, hypothesis driven experiments

Motivation and Aim: Modern proteomic techniques enable profiling of changes in levels of hundreds and thousands proteins simultaneously. It helps define "main players" of studying process, but do not aid in shaping possible molecular mechanisms underlying the process. The aim of this work was to develop the approach for building molecular network of cell response to experimental treatment: this approach should be based on proteins profiling data, and use the potential of continuously updated data on binary protein-protein interaction (PPI) on public domain.

Methods and Algorithms: As an initial data we used results of experiment for studying response of crayfish nerve tissue proteome to oxidative stress, caused by photodynamic treatment. The initial set of proteins we used was the list of proteins, which had demonstrated more than 25% level changes in response to photodynamic treatment. To bind proteins of the initial set into hypothetic network of PPI we used two approaches for defining possible binary protein-protein interactions: text mining of paper abstracts from PubMed (using Pathway Studio software from Ariadne Genomics Inc.); 2) querying PPI databases (mostly BioGrid DB, using Cytoscape platform).

Results: Both approaches we used enabled fast recognition of partners for proteins of the initial set, by analyzing huge arrays of data, which is backbreaking for manual analysis. The size of obtained networks could be adjusted. Originally, obtained PPI networks did not take into account the context of the experimental treatment – the networks included all met PPI facts. To take into account the process we study we filtered entities of the initial network using annotations (ontology) from Gene Ontology database (GO). We was lucky, and GO already had the ontology "response to oxidative stress" – which is an essence process of photodynamic treatment. The final network made possible to find out unknown in the context of photodynamic treatment interaction between two protein kinases, which are both hubs (nodes with many edges).

Conclusion: Our approach helps to infer hypotheses on protein-protein interactions underlying studying biological process, using all known PPI facts from different experimental models. Experimental verification of such hypotheses at new models would promote more system consideration the question about versatility and relevance of some newly-founded PPI to certain biological process. Generally, this could make faster and more reasonable adding new edges and entities to the "canonical" signaling and metabolic networks.

Availability: Pathway Studio demo version is available from Ariadne Genomics Inc. on-demand. Cytoscape is an open source platform and available at: www.cytoscape.org. PubMed, BioGrid and GO are public domain data bases.

THE SOLUTION OF THE INVERSE LIGHT-SCATTERING PROBLEM FOR PRECISE MORPHOLOGICAL CHARACTERIZATION OF MILK FAT GLOBULES

Konokhova A.I.*, Gilev K.V., Strokotov D.I., Yurkin M.A.,Maltsev V.P.
Institute of Chemical Kinetics and Combustion SB RAS, Novosibirsk, Russia
e-mail: konokhova_a@mail.ru
**Corresponding author*

Key words: inverse light-scattering problem, global optimization, milk fat globules

Motivation and Aim: Scanning flow cytometry proved itself as a method able to characterize single particles of different shapes by their morphology from light scattering [1]. The method is based on measurement of angle-resolved light-scattering patterns (LSPs) of individual particles and on the solution of the inverse light-scattering (ILS) problem. The ILS problem in its turn requires theoretical simulation of light scattering by a particle described by a proper optical model. The latter is a crucial factor for precise characterization.

Morphological characterization of milk fat globules (MFGs) is an important problem due to their function as a major carrier of bioactive molecules in milk. Thus, their characteristics play an important role in manufacturing of various dairy products. Usually modeled by spheres they nevertheless have a weak asphericity and can also be modeled as oblate spheroids. Our goal is to test different models for MFGs and develop an enhanced characterization method of the MFG population by their sizes and refractive indices.

Methods and Algorithm: LSPs of individual MFGs were measured with the Scanning Flow Cytometer. We used T-matrix method to simulate LSPs for the model of an oblate spheroid, and the Mie theory – for a homogeneous sphere. The ILS problem was solved through the global minimization and was performed by using DiRect algorithm for spherical model and nearest-neighbor interpolation with precalculated database of LSPs for spheroidal one. Both models were applied to each MFG, and the best one was chosen by applying the F-test.

Results: The inversion algorithm allowed us to separate measured MFGs into two fractions: those well-modeled by a sphere and those requiring a spheroidal model. While the spherical model was sufficient for large part of MFGs, spheroidal model was still required for the largest MFGs. This separation improved the accuracy of particle characterization over the whole sample in comparison with the results based on a single model.

Conclusion: We have demonstrated an advanced performance of the scanning flow cytometry in characterization of individual particles from light scattering. For the first time the shape of MFGs was modeled by an oblate spheroid to agree with sensitive measurement of angle-resolved light scattering. The solution of the ILS problem for individual MFGs allowed us to construct the distributions over MFG characteristics of the milk samples. This precise quantitative characterization of milk samples can be used to control manufacturing processes.

References:

1. V. Maltsev, A. Chernyshev, D. Strokotov. (2013) Light-Scattering Flow Cytometry: Advanced Characterization of Individual Particle Morphology. In: *Flow Cytometry: Principles, Methodology and Applications,* S. Papandreou (Ed.), 79–103 (Nova Science Publishers).

USER INTERFACES FOR MATHEMATICAL MODELS, BASED ON HUMAN ANATOMY

Kramarenko V.K.[1]*, Churikov A.Yu.[2], Zavodaev P.V.[1]
[1]*Moscow Institute of Physics and Technology, Moscow, Russia;*
[2]*Immanuel Kant Baltic Federal University, Kaliningrad, Russia*
e-mail: kramarenko.vasiliy@gmail.com
**Corresponding author*

Key words: mathematical modeling, natural user interface, human anatomy, bioimpedance, haemodynamics

Motivation and Aim: Today complexity of mathematical models is extremely increasing. Complexities of results and init data are also increasing. These trends lead to creating new systems for interaction between human and mathematical code. Evolution of user interfaces from punch cards to head-displays can help us to make such system. Due to these facts we created two interfaces, that enable us to interact with mathematical models.

Methods and Algorithms: Present interfaces are based on two different methods. First technology is connected with large multitouch panels. These panels use technology, which differs from tablet devise because of large size of screen. Our panel detects touch by analyzing finger shadows. The other important device is connected with second interface and represents new age of user interfaces. It is Oculus Rift, the gadget for creation of complemented reality.

First interface is created for two mathematical models and, consequently, uses two different programs. The first program is Paraview [1]. This program, initially, is created for dealing with large variety of meshes, in particular different results of simulations. It is used for analysing results of bioimpedanse measurements [2]. Other program is Blender [3]. This program has more universal tools due to its function - creation of 3D scenes and cartoons. Based on this program interface was created for haemodynamic model [4].

The other interface is based on head-display Oculus rift [5] and Leap Motion [6] sensor, that helps to create visual reality and opens up new perspectives for control of mathematical models. Oculus Rift is a system, which contains 2 displays and system of giro, that gave us data about head rotation. Leap motion is a gadget which contains sources and detectors of light. This devise help us to detect finger motion and gestures. The program base of our new interface is Unity [7], that already have special procedures for connection with Oculus and Leap motion.

Results: Two interfaces were created for simplification of mathematical models using The foundation of more complex interface was laid.

Conclusion: Natural user interfaces for mathematical models were created. Such programs can simplify the using of mathematical models and became the fundament of on-line patient-orientated system.

Availability: Now interfaces are not presented in the internet, because of developing.

References:

1. A. Cedilnik, B. Geveci, K. Moreland, J. Ahrens, and J. Favre, Remote Large DataVisualization in the ParaView Framework. A. Heirich, B. Raffin, L. P. Santos (eds.) editors, Eurographics Parallel Graphics and Visualization 2006, pages 162-170, May 2006.
2. Yu. Vassilevskii, S. Simakov, V. Salamatova, Yu. Ivanov, T. Dobroserdova. Numerical issues of modelling blood flow in networks of vessels with pathologies. Russian Journal of Numerical Analysis and Mathematical Modelling, Vol. 26, No. 6, 2011, pp. 605-622.
3. www.blender.org
4. A.A. Danilov, V.K. Kramarenko, D.V. Nikolaev, S.G. Rudnev, V.Yu. Salamatova, A.V. Smirnov, Yu.V. Vassilevski Sensitivity field distributions for segmental bioelectrical impedance analysis based on real human anatomy // J. Phys.: Conf. Series, 2013, 434: 012001DOI: 10.1088/1742-6596/434/1/012001
5. http://www.oculusvr.com/
6. F. Weichert, D. Bachmann, B. Rudak; D. Fisseler, Analysis of the Accuracy and Robustness of the Leap Motion Controller. Sensors 2013, 13, 6380-6393.
7. https://unity3d.com/

ROTATIONAL DYNAMICS OF BASES IN THE GENE CODING INTERFERON ALPHA 17 (IFNA17)

Krasnobaeva L.A.[1, 2]*, Yakushevich L.V.[3]
[1]*Siberian State Medical University, Tomsk, Russia;*
[2]*Tomsk State University, Tomsk, Russia;*
[3]*Institute of Cell Biophysics RAS, Pushchino, Moscow Region, Russia*
e-mail: kla1983@mail.ru
**Corresponding author*

Key words: interferon alpha 17 (IFNA17), rotational oscillations of bases, DNA

Motivation and Aim: The group of proteins – interferons – plays an important role in medicine. Due to them cells become immune to viruses. Therefore the studying of the genes coding these proteins is an interesting and actual task. In the present work, rotational oscillations of nitrogenous bases in the sequence of the gene coding interferon alpha 17 (IFNA17), are investigated.

Methods and Algorithms: As a mathematical model simulating oscillations of the bases, we use a system of two coupled nonlinear partial differential equations that take into account the effects of dissipation, the action of external fields and the dependence of the coefficients on the sequence of bases. We apply the methods of the theory of oscillations to to solve the equations in linear approach and to construct dispersion curves. To consider the system of the equations in the general (nonlinear) case, the approximation of the average field, which allows to reduce the problem of two coupled equations to the problem of a single equation simulating oscillations of the bases of one of two polynucleotide chains in the average field, induced by the second polynucleotide chain is was used. This equation has been solved by two methods: the method of the concentration [1], and the energy method [2].

Results: In the linear approach, the solutions of the model system of differential equations have been obtained, and the dispersive curves determining the dependence of the frequency of the plane waves (ω) on the wave vector (q) have been constructed. In the nonlinear case, the solution in the form of was obtained, and the main characteristics of the kink: the energy density (ρ), the total energy (E), the rest energy (E_0), the rest mass (m_0) and the size (d), were calculated. With the help of energy method the kink velocity (υ), the path (S) that kink passed, and the lifetime (τ) have been obtained.

References:

1. L.V. Yakushevich, L.V. Krasnobaeva. (2008) A new approach to studies of non-linear dynamics of kinks activated in inhomogeneous polynucleotide chains, *International Journal of Nonlinear Mechanics*, **43:** № 10, 1074–1081.
2. D.W. McLaughlin, A.C. Scott. (1978) Perturbation analysis of fluxon dynamics, *Phys. Rev. A*, **18:** 1652-1660.

POWER LAW FOR RANK DISTRIBUTION OF GENE DENSITY IN HUMAN GENOME PROJECTS

Lenskiy S.V.*, Lenskaya T.I.
Ural Federal University, Ekaterinburg, Russia
e-mail: sergeylensky@rambler.ru
**Corresponding author*

Key words: *gene* density, rank distribution, power law, human genome

Motivation and Aim: Rank distribution is useful for investigating nucleotide sequences [1, 2]. Application of rank distribution to complementary pairs gave an opportunity to find a pattern for assessment of order in eukaryotic genomes of different origin, to construct theoretical model of the most statistically ordered genome and to identify the real representative with these properties – Arabidopsis thaliana [3]. In this research rank distribution was used for studying the density of genes in the human genome.

Methods and Algorithms: Now three human genome projects located on NCBI site (http://www.ncbi.nlm.nih.gov/genome/) are suitable for estimating the density of genes on autosomes: The Genome Reference Consortium project, the project of J. Craig Venter Institute project and the project of Washington University School of Medicine. The gene density of an autosome is the ratio of a number of genes located on the autosome per size of the autosome in million base pairs (Mb).The average gene density values of autosomes were arranged in rank order. We used the least squares method to find an approximation curve for the rank distribution. The coefficient of determination (R^2) was used for measuring of how well the approximation had predicted data of the genome projects.

Results: It was found that the rank distribution of average gene density of autosomes in the human genome can be well approximated by a power function: $f(x) = a*x^b$, where a and b were parameters of an approximation curve. Good approximation was achieved by a equals to the highest value of average gene density values in the human genome and b = -0.5. Based on found parameters we constructed the model for the rank distribution of gene density. The coefficient of determination (R^2) for the human genome projects were in the range $0.96 < R^2 < 0.98$. Approximation of genome data from different sources with high values of coefficient of determination indicated a presence of a close connection between the power law and the data. We had observed the rank distribution of gene density obeyed a power law. We found good approximation of genetic load on autosomes in the human genome.

Conclusion: The found approximation of gene density distribution is probably related to an allocation of functional load to autosomes in the human genome. The revealed regularity points to difference in functional role of autosomes in the genomes. Further it'll help to clarify different biological roles of autosomes in the genome. The rank distribution can help to compare and unify genome research projects. Rank research tools can be and ought to be constantly enhanced.

References:

1. R.N. Mantegna et al. (1994) Linguistic Features of Non-coding DNA Sequences, *Phys. Rev. Lett.,* **73**: 3169-3172.
2. P.A. Tsonis, A.A. Tsonis. (2002) Linguistic features in eukaryotic genomes, *Complexity*, **7(4):** 13-15.
3. S.V. Lenskiy, T.I. Lesnova. (2007) Searching of Statistically Ordered Eukaryote Genome, *ISMB/ECCB 2007*, poster session (O18), 21–25 July, Vienna, Austria.

ON A MODIFICATION OF THE THEORETICAL BASIS OF THE PENROSE-HAMEROFF MODEL OF CONSCIOUSNESS

Levichev A.[1]*, Palyanov A.[2]
[1]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;*
[2]*A.P. Ershov Institute of Informatics Systems SB RAS, Novosibirsk, Russia*
e-mail: alevichev@gmail.com
**Corresponding author*

Key words: Penrose-Hameroff model of consciousness, Orchestrated objective reduction (Orch-OR), *C. elegans* nematode, DLF-theory, separation between space-times

Theoretical physicist R. Penrose and anesthesiologist S. Hameroff proposed in the mid 1990's that consciousness depends on biologically 'orchestrated' coherent quantum processes in collections of microtubules within brain neurons. Recently they reviewed Orch OR in light of criticism and developments in quantum biology, neuroscience, physics [1, 2]. They theorize that the wave function cannot be sustained in superposition beyond a certain energy difference between the quantum states. Technically, such methods are based on introducing of a Newtonian limit in General Relativity (GR). They define the notion of separation between space-times and give an approximate value for the energy difference. As an example of an application of this estimate, the nematode worm *C. elegans* would require one third of its tubulins to sustain quantum coherent superposition for 500 msec. Penrose and Hameroff consider it unlikely, but not altogether impossible. Our continuing experience in computational simulation of this organism [3] hopefully will be augmented by investigations in this field.

The goal of our study is to modify the Penrose-Hameroff model by using the DLF-theory [4, 5] instead and by introducing an alternative notion of separation between space-times.

We introduce a certain class K of GR homogeneous space-times (or worlds). Within the *compact model* [4], each of these worlds is a dense subset of the Lie group U(2) – see [4, Theorems 6 and 10, and 5, Theorem 2], where the worlds D, L, F are introduced and described. The entire class K is obtained by applying (conformal) transformations g given by [4, formula (1.1)]. These transformations form a matrix group SU(2,2) which is one of the fundamental symmetry groups of modern physics. To introduce an alternative definition of the separation between space-times, we use the structural theory of Lie groups and methods of pseudo-Riemannian differential geometry. It is instrumental that the group U(2) admits a positive definite bi-invariant metric (besides the Lorentzian one). Our approach resulted in two possible definitions of separation between worlds from the class K. One is based on the Iwasawa decomposition in SU(2,2). The second possibility is based on the integration of the conformal factor of the (above introduced) transformation g which maps one world from K onto another. The integration is carried out separately over each region where the conformal factor is greater than 1 (or less than one). Each time the integral is compared to the volume of such a region.

Conclusion: We have suggested a modification of the theoretical basis of the Penrose-Hameroff model of consciousness by introducing a new class K of space-times. It is important that any two worlds from K have isomorphic causal structures. The modification resulted in two possible definitions of separation between space-times. Preliminary, we argue for the one which is based on 4D integration of the conformal factor. It is expected that results of our study will help to clarify the question of validity of the original Penrose-Hameroff model of consciousness.

References:

1. S. Hameroff, R. Penrose. (2014) Reply to criticism of the 'Orch OR qubit' – 'Orchestrated objective reduction' is scientifically justified, Physics of Life Reviews, 11 (1): 94–100.
2. S. Hameroff, R. Penrose. (2014) Consciousness in the universe: A review of the 'Orch OR' theory, Physics of Life Reviews (Elsevier), 11 (1): 39-78.
3. M. Vella, A. Palyanov, P. Gleeson, S. Khayrulin. (2013) Integration of predictive-corrective incompressible SPH and Hodgkin-Huxley based models in the OpenWorm in silico model of C. elegans. BMC Neuroscience, 14(Suppl. 1): P209.
4. A.V. Levichev. (2011) Pseudo-Hermitian realization of the Minkowski world through DLF theory, Physica Scripta, 83 (1): 1-9
5. A.V. Levichev, J. Feng. (2013) More on the Mathematics of the DLF Theory: Embedding of the Oscillator World L into Segal's Compact Cosmos D, AJUR, 11 (3-4): 29-33.

RECONSTRUCTION OF THE MOUSE BRAIN VASCULAR NET ACCORDING TO THE DATA OF HIGH-FIELD MRI-SCANNER

Maltseva S.V.[1, 2]*, Akulov A.E.[3], Derevtsov E.Yu.[1, 2], Cherevko A.A.[4, 2], Chupakhin A.P.[4, 2], Khe A.K.[4, 2]

[1]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[4]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia*
e-mail: sv_maltseva@mail.ru
**Corresponding author*

Key words: MRI, brain vasculature modelling, approximation, noise reduce, segmentation

Motivation and Aim: the problem is to reconstruct mouse brain vasculature [1] by given tomographic data; the recovery of the vascular net should be one-connected. This problem is very important not only for fundamental sciences but also for application once (medicine, biology).

Methods and Algorithms: the reconstruction is realized by automatic segmentation according to the mathematically preprocessed data. Initial data are some tomographic packets with the different spatial slice orientation. Preprocessed data are obtained from data packets using some additional technics allowing to correct data intensity, to make vessels walls more contrast and to decrease noise.

Results: the algorithm proposed allows to reconstruct the vascular net more efficient as compared with standard tomographic reconstruction.

Conclusion: detailed and one-connected geometrical model of the vascular net allows to realize haemodynamic CFD-modelling [2] and to discover correlations between structural features of the net and physiological characteristics [3] of the animal units.

Availability: not available.

Acknowledgements: RFBR 14-01-00036, RFBR 14-04-31829-mol_a, Interdisciplinary integration project SB RAS 60 and 61, SB RAS IP № 44, Department of Mechanics RAS № 2.13.4.

References:

1. S.W. Oh et al. (2014) A mesoscale connectome of the mouse brain, Nature, 508, 207-214.
2. C.G. Caro et al. (1979) The mechanics of the circulation, Cambridge University Press.
3. A.D. Nozdrachev, E.L. Polyakov. (2001) Anatomy of the rat, St. Petersburg, Lan'.

PROBABILISTIC FORMAL CONCEPTS WITH NEGATION

Martynovich V.V.[1]*, Vityaev E.E.[2]
[1] *Novosibirsk State University, Novosibirsk, Russia;*
[2] *Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia*
* *Corresponding author*

Key words: formal concepts, notions, data mining, association rules, classification

Motivation and aim: The paper tends to generalize somewhat famous approach brought by Formal Concept Analysis. The main improvement related with including the negations into concept intent construction. Also statistical ambiguity problem should have been eliminated by consistency and soundness results.

Methods and algorithms: The long ways starts from classic Formal Concept Analysis structures and tools. Previous works stated the definition of formal concept as fixed points for prediction operator. So the introduction of probability on formal system leads to key notion of probabilistic concept. Using the ideas of semantic probabilistic inference, we propose some results about such concepts.

Results: The formal concepts with negations were introduced, prediction operator properties has been researched, and fixed point consistency and soundness theorems has been proven. Additionally, the inconsistency case has been studied and described. Finally, those lead to algorithm, which yields a list of all probabilistic concepts.

Conclusion: Algorithm in paper could be significantly improved. Broad theoretical field of probabilistic formal context is a subject of subsequent researches. Presented method deserves more attention and will be developed into independent Data Mining methods.

Availability: Full text is not available online right now, but may be provided by personal request via emailing.

References:

1. Ganter, B.: Formal Concept Analysis: Methods, and Applications in Computer Science. TU Dresden (2003).
2. Ganter, B., Wille, R.: Formal concept analysis. Mathematical Foundations. Berlin-Heidelberg-New York, Springer (1999).
3. Vityaev, E.E., Demin, A.V., Ponomaryov, D.K.: Probabilistic Generalization of Formal Concepts. Programming and Computer Software. 38(5), 219-230 (2012).
4. Missaoui, R., Kwuida, L.: Implications in Triadic Formal Contexts. In: 9th International Conference, ICFCA 2011. Nicosia, Cyprus, Springer (2011).
5. Vityaev, E., Kovalerchuk, B.: Empirical Theories Discovery based on the Measurement Theory. Mind and Machine 14(4), 551-573 (2004).
6. Vityaev, E.E.: Knowledge discovery. Computational cognition. Cognitive process models. Novosibirsk State University Press, Novosibirsk (2006) (in Russian).
7. Vityaev, E., Smerdov, S.: On the Problem of Prediction In: KONT/KPP 2007, LNAI 6581, pp. 280-296, Springer, Heidelberg, (2011).
8. Speransky, S.O.: Logic of probability and probability logic. Novosibirsk State University, Novosibirsk, PhD thesis. (2013) (in Russian).
9. Vityaev, E.E., Neupokoev, N.V.: Formal model of perception and pattern as fix point of anticipations. In: Approaches to the thinking modeling, pp.155-172, Moscow, URSS Editorial (2014).

APPLICATION OF BAYESIAN METHODS TO THE PROBLEM OF CLASSIFICATION OF NATURAL LANGUAGE TEXTS

Maslov P.P.*, Berikov V.B.
Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia
e-mail: berikov@math.nsc.ru
**Corresponding author*

Key words: classification, clustering, natural language processing, Bayesian models

Motivation and Aim: Information on a particular theme, as a rule, is represented by a large volume of documents. Correct thematic division of documents on similar groups can greatly increase the efficiency of their use. For this reason the problem of classification of text documents is very important.

Methods and Algorithms: The formal-declarative natural language processing method used Bayesian models on weighted sets of semantic nodes production rules (chains), extracted from a set of text documents is suggested. Text processing includes the following steps: graphematic, morphological, syntactic, shallow semantic [1]. While processing, each text document is represented by the model: $<D,S,A,M,R,T>$ where: D is the text document, S is the set of text objects, R is the set of relations among objects and T is the set of syntactic and shallow semantic rules for binding objects S. The final step in the text processing is the classification of documents by Bayesian model on a finite set of events [2]. Assigning the document to a particular class comes to the determination of the conditional probability $P(F|C) = P(Fs_i|C) * P(Fr_j|C)$, where F is determined by two variables: Fs_i, Fr_j - frequency of occurrence of objects and relations between them. The probability that D is belong to class C is defined by expressions: $P(C|D) = P(D \cap C)/P(D)$, $P(D|C) = P(D \cap C)/P(C)$ and $P(C|D) = P(D|C) * P(C)/P(D)$. For n classes the probability of assigning the document to the *K-th* class is defined as $P(D|C_k) = \prod_i p(Fs_i|C_k) * \prod_j p(Fr_i|C_k)$, $k = 1..n$.

Results: Algorithms and software tool allowing analysis and classification of texts are developed.

Conclusion: The extracted objects and relations can be used for documents clustering. The degree of proximity between texts is calculated by the occurrence frequency, using maximum likelihood method. To improve the stability of solutions, the ensemble approach is applied [3].

References:

1. Maslov P.P. Detecting and extracting the causal patterns of text in natural language // Materials of all-Russian conference with international participation “Knowledge-Ontology-Theory” (ZONT-09), Volume 2. P. 3-10.
2. Berikov V.B., Lbov G.S. Bayes estimates for recognition quality on finite sets of events // Doklady Mathematics. - 2005. - Vol. 71, № 3. - P. 327-330.
3. Berikov V. Weighted ensemble of algorithms for complex data clustering // Pattern Recognition Letters. 2014. Vol. 38. P. 99-106.

EQUATION OF STATE OF BLOOD FLOWS IN SMALL VESSELS

Medvedev A.E.
Khristianovich Institute of Theoretical and Applied Mechanics SB RAS, Novosibirsk, Russia
e-mail: medvedev@itam.nsc.ru

Key words: hematocrit, Poiseuille flow, the mathematical model, two-phase flow, the relative viscosity

Motivation and Aim: The blood flow in small vessels has some specific features observed in vitro and in vivo [1]: (I) Fahraeus effect, which is the hematocrit dependence on the vessel diameter; (II) existence of a cell-depleted (cell-free) layer of the plasma near the wall; (III) less steep velocity profile (as compared with the Poiseuille flow profile); (IV) Fahraeus-Lindqvist effect, which is an explicit dependence of blood viscosity on the blood vessel diameter. Available rheological models of blood fail to capture this effect. Therefore, a two-phase model of the blood flow in small vessels is proposed.

Methods and Algorithms: Blood is considered as a two-phase medium consisting of the plasma (viscous incompressible fluid) and blood elements (mainly erythrocytes) ([2]).

Results: Fairly simple unified formulas are obtained to describe the blood motion in vessels with diameters ranging more than 10 μm. This model explains why the blood viscosity decreases with decreasing vessel size, why the hematocrit depends on the vessel diameter, why the velocity profile differs from the Poiseuille flow profile, and why the near-wall plasma layer is formed.

Conclusion: A two-phase blood model for describing the blood flows through the vessels is presented. The generalization of the Poiseuille solution is obtained for the flow with the viscosity coefficient varying over the vessel cross section; the velocity profile is less steep for this flow than that for the Poiseuille flow. The model describes the known features of the blood flow in vessels: the dependence of the hematocrit level on the vessel diameter, the existence of the near-wall plasma layer, a less steep blood velocity profile (in comparison with the Poiseuille flow profile), and the dependence of the blood viscosity on the vessel diameter. Analytical dependences for the blood velocity and viscosity as functions of the blood vessel diameter and hematocrit level were obtained.

Availability: A unified two-phase model for describing blood flows in small vessels is proposed in this paper. Based on this model, a mathematical explanation is given for effects of blood flows in vessels, which have been known for a long time [1].

References:

1. A.R. Pries, T.W. Secomb. (2008) Blood Flow in Microvascular Networks, In: *Handbook of Physiology: Microcirculation.* Tuma R.F. et al. (Eds.), 3-36 (Academ Press).
2. A.E. Medvedev, V.M. Fomin. V.I. Samsonov. (2012) Mathematical modeling of the blood flow in blood vessels, In: *Circulatory System and Arterial Hypertension: Experimental Investigation, Mathematical and Computer Simulation*. L.N. Ivanova et al. (Eds.), 55-114 (New York: Nova Science Publishers, Inc.).

COMMUNITY STRUCTURE OF WEB-GRAPHS OF ACADEMIC INSTITUTIONS

Medvedev A.N.
Institute of Mathematics SB RAS, Novosibirsk, Russia
e-mail: an_medvedev@yahoo.com

Key words: network science, web graph, community detection, academic networks

Motivation and Aim: web-graph is a network of web pages, represented as nodes, and links between them, represented as edges. The analysis of academic networks naturally represented as graphs has received substantial interest in recent years [1, 5]. Nowadays there appear more and more evidence that real-life network structure is far from being homogeneous. The structural heterogeneity may be revealed by studying the community structure of the network. Community detection as now one of the hot topics in network science, received a significant development in a past decade [6]. In this work we introduce the datasets of academic web-graphs of Siberian Branch of Russian Academy of Sciences, Russia, and of Fraunhofer-Gesellschaft, Germany, and study their community structure.

Methods and Algorithms: we study these graphs using community detection algorithms based on maximization of a weighted modularity function Q for non-overlapping communities [7] and clique percolation method for overlapping ones [8]. The high precision heuristics are used for Q maximization, mainly the tabu search algorithm [9].

Results: our analysis shows that the networks are very centralized; however it is possible to study their community structure that reflects the collaboration and connection between organizations in these networks on the level of web-space.

Acknowledgements: this work has been supported by the Interdisciplinary Integration Project of SB RAS N.21 under the title «Investigation of regularities and trends of self-organizing systems on the examples of the Web space and biological communities», Grant 12-01-00448 of the Russian Foundation of Basic Research and Grant NSh-1939.2014.1 of President of Russia for Leading Scientic Schools.

References:

1. M. Thelwall, D. Wilkinson. (2003) Graph structure in three national academic webs: power laws with anomalies, Am. Soc. Inf. Sci. Technol., 54(8), 706–712.
2. D. Stuart, M. Thelwall, G. Harries. (2007) UK academic web links and collaboration — an exploratory study, J. Inf. Sci., 33(2), 231–246.
3. А.А. Печников. (2013) Применение вебометрических методов для исследования информационного веб-пространства научной организации (на примере Карельского научного центра РАН), Труды КарНЦ РАН. No 1. Сер. Математическое моделирование и информационные технологии, 4, 86–95.
4. А.Ю. Веснин, Е.В. Константинова, М.Ю. Савин. (2013) О сценариях присоединения новых сайтов к веб-пространству СО РАН, Вестник НГУ, серия: информационные технологии. Т. 11. № 4, 28–37.
5. Yu.I.Shokin,A.Yu.Vesnin,A.A.Dobrynin,O.A.Klimenko,E.V.Konstantinova,I.S.Petrov,E.V.Rychkova. (2014) Investigation of the academic web space of the Republic of Serbia, Zbornik radova konferencije MIT 2013, Belgrad, Serbia, 601–607.
6. S. Fortunato. (2009) Community detection in graphs. CoRR, abs/0906.0612.
7. M.E.J. Newman. (2004) Analysis of weighted networks. Phys. Rev. E, 70(5):056131.
8. G. Palla, I. Derényi, I. Farkas, T. Vicsek. (2005) Uncovering the overlapping community structure of complex networks in nature and society, Nature 435, 814.
9. A. Arenas, A. Fernández, S. Gómez. (2008) Analysis of the structure of complex networks at different resolution levels, New Journal of Physics, 10, 053039.

SIBERIAN SUPERCOMPUTER CENTER AS A SERVICE FOR BIOINFORMATICS

Mikhailenko B.G.[1], Glinskiy B.M.[1], Kuchin N.V.[1]*, Chernykh I.G.[1]
[1]*Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia*
e-mail: kuchin@sscc.ru
**Corresponding Author*

Key words: supercomputers, bioinformatics, high performance computing

Motivation and aim: The Siberian Supercomputer Center (SSCC) was founded at the Institute of Computational Mathematics and Mathematical Geophysics (ICM&MG) SB RAS in 2001. The SSCC main objectives are development and using the supercomputer technologies for the numerical simulation, providing high performance computer resources to researchers from the institutes of SB RAS and Universities. A special attention is being given to training specialists of the SB RAS and university students in parallel scalable computing methods and algorithms on supercomputers as well as in methods of solving large-size problems.

Results: Currently there are two clusters in the SSCC that are exploited in the multi-access mode. One of the two clusters is based on the Intel Xeon processor (the MPP-architecture), its peak performance is 30TFlops, the MPI- and OpenMP-aided programming. The other cluster NKS-30T has a hybrid architecture with the NVIDIA Tesla M2090 GPU (the GPGPU-architecture). It's peak performance is 85 TFlops, parallel programming in C/C++ CUDA and OpenCL. The seminars "Architecture, System and Application Software of Cluster Supercomputers" are regularly held on the basis of the SSCC, the Chair of Computer Systems of NSU and the Competence Center on High-Performance Computing of SB RAS - Intel.

Conclusion: Especially for bioinformatics tasks SSCC is equipped with HP DL980 G7 Server with 8x Intel Xeon E7-4870 CPUs (80 cores) and 1TB RAM. This configuration is reliable for big data problems such as genome sequencing and assembling, molecular dynamics simulation. SSCC provide an access for GROMACS open source package for molecular dynamics problem simulation. GROMACS is well designed for NKS-30T hybrid architecture.

Availability: http://www2.sscc.ru/, http://www2.sscc.ru/HKC-30T/HKC-30T.htm

Acknowledgements: This work was partially supported by RFBR grants No. 13-07-00589, 14-01-31199.

CONSTRUCTION OF THE HEMODYNAMIC MODEL BASED ON CLINICAL DATA: INVERSE PROBLEM

Mikhaylova A.V.[1], Cherevko A.A.[1, 2], Chupakhin A.P.[1, 2]*, Krivoshapkin A.L.[3], Orlov K.U.[3]

[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia;*
[3] *Meshalkin Novosibirsk State Research Institute of Circulation Pathology, Novosibirsk, Russia*
e-mail: alexander190513@gmail.com
**Corresponding author*

Key words: modeling, hemodynamic, inverse problem, nonlinear oscillator, stable cycles

Hemodynamic modeling of cerebral vessels is a complex and important issue that has both fundamental and evident medical applications. Desired model should describe an unsteady, pulsatile blood flow in a complex net of vessels reacting on the flow pulsations. Such a model should be based on experimental data.

In the Novosibirsk Institute of Circulation Pathology jointly with the Institute of Hydrodynamics the measurements of pressure and velocity of blood flow in the vessels of the human brain in the presence of anomalies, such as arteriovenous malformations, are carried out. Using these experimental data and methods of the theory of inverse problems, a mathematical model - a second order differential equation of a nonlinear oscillator type with the right-hand side – is constructed:

$$q''+c_1q'+c_2q+c_3q^2+c_4q'q^2+c_5q^3=c_6u\,. \qquad (1)$$

Here, blood flow velocity u is a control, q is pressure in the blood vessel. The unknown quantities in the inverse problem are the coefficients of equation (1). Coefficients c_2, c_3, c_5 describe elastic properties of blood vessels, and c_1, c_4 correspond to viscous friction.

This equation describes periodic oscillations of pressure and velocity at a fixed state of the patient (sick or healthy). In terms of solutions of a nonlinear oscillator model this process is described by a closed curve (limit cycle) on the “velocity-pressure” plane.

The proposed method allows us to calculate the pressure for a given velocity on large intervals of time (about 10 minutes), by constructing a differential equation using a small period of time (a few seconds). The resulting model approximates clinical data with good accuracy. Obtained equation has structural stability and the stability to the initial data.

This work was financially supported by SB RAS, FIP number 44; OEMMPU project № 2.13.4; RFBR project № 14-01-00036.

IDENTIFICATION OF BASED ON EXPERIMENTAL CLINICAL DATA HEMODYNAMIC MODEL

Mikhaylova A.V.[1]*, Cherevko A.A.[1, 2], Chupakhin A.P.[1, 2], Krivoshapkin A.L.[3], Orlov K.U.[3]
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia;*
[3]*Meshalkin Novosibirsk State Research Institute of Circulation Pathology, Novosibirsk, Russia*
e-mail: little_hedgehog@bk.ru

Key words: system identification, hemodynamic, clinical data

Motivation and Aim: Hemodynamic modeling of cerebral vessels is a complex and important issue that has both fundamental and evident medical applications. Desired model should describe an unsteady, pulsate blood flow in a complex net of vessels reacting on the flow pulsations. Such a model should be based on experimental data.

Methods and Algorithms: In the Novosibirsk Institute of Circulation Pathology jointly with the Institute of Hydrodynamics the measurements of pressure and velocity of blood flow in the vessels of the human brain in the presence of anomalies, such as arteriovenous malformations, are carried out. Using these experimental data and methods of the theory of inverse problems, a mathematical model - a second order differential equation of a nonlinear oscillator type with the right-hand side – is constructed. This equation describes periodic oscillations of pressure and velocity at a fixed state of the patient (sick or healthy). In terms of solutions of a nonlinear oscillator model this process is described by a closed curve (limit cycle) on the "velocity-pressure" plane.

Results: The proposed method allows us to calculate the pressure for a given velocity on large intervals of time (about 10 minutes), by constructing a differential equation using a small period of time (a few seconds). The resulting model approximates clinical data with good accuracy.

Conclusion: The results can be used for validation of more complicated hemodynamic models and determine the state of the vascular bed.

Availability: not available.

Acknowledgements: RFBR 14-01-00036, SB RAS IP № 44, RAS department of Mechanics № 2.13.4.

CHARACTERIZATION OF BLOOD PLATELETS SOLVING THE INVERSE LIGHT-SCATTERING PROBLEM WITH PRE-COMPUTED INTERPOLATING SET

Moskalensky A.E.*, Strokotov D.I., Yurkin M.A., Maltsev V.P.
Institute of Chemical Kinetics and Combustion SB RAS, Novosibirsk, Russia;
Novosibirsk State University, Novosibirsk, Russia
e-mail: sunmosk@mail.ru
**Corresponding author*

Key words: inverse light-scattering problem, scanning flow cytometer, light-scattering pattern, blood platelets, platelet activation

Motivation and aim: The scanning flow cytometer [1] allows one to measure light-scattering patterns (LSPs) of individual biological cells, i.e., the intensity of scattered light *versus* the polar scattering angle. The characterization of cells, described by a few parameters, from the measured LSPs requires the solution of inverse light-scattering problem by global optimization. For particles of complex shape, direct fitting is unfeasible due to large time of solution of direct problem. The aim of this research is to develop a method of solution of inverse light-scattering problem with pre-calculated interpolating set, including error estimation.

Methods and algorithms: We used scanning flow cytometer for measurement of LSPs of individual human blood platelets. The geometry of platelets was approximated by oblate spheroid. We used discrete-dipole approximation (DDA) [2] for the solution of direct light-scattering problem for 500.000 spheroids with different parameters corresponding to blood platelets. All calculations were performed at the supercomputing center of Novosibirsk State University [3]. Then we used them as an interpolating set for the global optimization, determining parameters of each measured blood platelets. The Bayesian approach was used to estimate errors of each parameter.

Results: The method was used for characterization of blood platelets. The accuracy of the solution of inverse light scattering problem was quite good, resulting in sub-diffraction precision in the majority of cases. The determined parameters of blood platelets showed good agreement with literature data. The shape of platelets was changed during activation in accordance with well-known effect of platelets shape change.

Conclusion: The solution of inverse light-scattering problem using pre-computed interpolating set gives accurate results. This method was used for characterization of blood platelets and showed good agreement with literature. It is the only feasible approach for particles with complex geometries.

References:

1. Maltsev V.P. Scanning flow cytometry for individual particle analysis. Rev Sci Instrum 2000;71:243–55.
2. Yurkin M.A., Hoekstra A.G. The discrete-dipole-approximation code ADDA: Capabilities and known limitations. J Quant Spectrosc Radiat Transf 2011;112:2234–47.
3. Supercomputing center of Novosibirsk State University, http://www.nusc.ru

HIGH PERFORMANCE COMPUTING SIMULATION OF EVOLUTIONARY PROCESSES IN BACTERIAL COMMUNITIES

Mustafin Z.S., Lashin S.A.*
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;
Novosibirsk State University, Novosibirsk, Russia
e-mail: mustafinzs@bionet.nsc.ru
**Corresponding author*

Key words: high performance computing, parallelization, CUDA, MPI, OpenMP, evolution, simulation, mathematical model, multiscale model, software

Motivation and Aim: The simulation of the evolution of bacterial communities is an actual problem of modern computational biology. Haploid Evolutionary Constructor (HEC) is a software package which is used for the simulation of evolution of prokaryotic communities. HEC considers the following of biological organization: metabolic, genetic, population and ecological. Simulation of bacterial communities of extremely high genetic diversity can take dozens of hours. In this work we solve this problem with various high performance algorithms (for CPU and GPU, for distributed-memory and shared-memory systems).

Methods and Algorithms: We have developed 3 high population reproduction function, which is the most time-consuming one in HEC. The first version was implemented using OpenMP technology for shared-memory systems. The second was implemented using MPI technology for distributed-memory systems. The third version was developed using CUDA (Compute Unified Device Architecture) for simulations with GPU. The tests and benchmarks have been performed on the computer clusters of the Center for collective usage "Bioinformatics" of the SB RAS (http://bioinformatics.bionet.nsc.ru/), and the Siberian Super Computer Center (http://www2.sscc.ru/). For shared-memory version testing and comparison, we used PC with AMD Phenom II X6 series CPU and Nvidia GTX 570 GPU).

Results: The time of simulation was decreased from hours to minutes. MPI and Open MP versions showed near-linear acceleration (for MPI version it was also obtained with using 256 cluster threads). CUDA version showed a great decreasing of time in simulations of high variety bacterial systems (time decreased from 200 second for one generation to 3 second). We created classification for best using of different high performance versions with different bacterial variety.

Conclusion: Developed algorithms allow user to simulate situations which are very difficult to simulate in vitro because of huge bacterial community variety and size. With the new algorithms, the user can increase biological variety of simulated systems without spending an additional time. New algorithms support simulations with using CPU, GPU, PC and high performance clusters.

Availability: http://evol-constructor.bionet.nsc.ru

Acknowledgements: The work is supported by the RFBR grant 13-04-00620, the Stipend of the President of Russian Federation, the Program of the RAS Presidium № 28 "Evolution and biosphere origin".

References:

1. Lashin S.A., Matushkin Yu.G. Haploid evolutionary constructor: new features and further challenges // In Silico. Biol. 2012. V.11. N. 3. P. 125-135.

BOINC-BASED DESKTOP GRID INFRASTRUCTURE FOR VIRTUAL DRUG SCREENING♣

Nikitina N.N.[1, 2]*, Ivashko E.E.[2], Gupta Y.[1], Ludwig R.[1], Möller S.[1]
[1]*University of Lübeck, Department of Dermatology, Ratzeburger Allee 160, 23538 Lübeck, Germany;*
[2]*Institute of Applied Mathematical Research, Karelian Center of the Russian Academy of Sciences, Petrozavodsk, Russia*
e-mail: nikitina@krc.karelia.ru
**Corresponding author*

Key words: virtual drug screening, protein-ligand docking, desktop grid, volunteer computing, BOINC

Motivation and Aim: Virtual drug screening aims at specifying and ranking a set of drug candidates according to their predicted binding properties against target proteins. The considerable computational effort typically prevents an adoption of the technology in small research environments. This work demonstrates how *in silico* screening can be facilitated by volunteer computing concepts to complement local research questions in preclinical wet-lab environments.

Methods and Algorithms: Large sets of compounds are freely available from the ZINC database {{http://zinc.docking.org}} and are docked using AutoDock Vina software {{http://autodock.scripps.edu}}. Computational load was distributed with the BOINC middleware {{http://boinc.berkeley.edu}}.

Results: The setup was implemented within the Lübeck Department of Dermatology with research emphasis on autoimmune blistering diseases. Over 40 desktop grid clients, locally and remotely as provided by friends and families of the researchers, have completed over four million computational jobs in two months. The compute facilities allowed to *in silico* confirm the docking scores for multiple structural variants of the receptor. Among many thousands of compounds with higher affinity than the known natural ligand are 20 of known drugs. The priorisation of the many will depend on the results of the *in vitro* validation of those already FDA-approved compounds. Further iterations between wet- and dry-labs are expected.

Conclusion: This work proves the readiness of the virtual screening technology in preclinical environments. Additional work is required to allow for IT skilled biochemists to adopt the technology for themselves and overcome hurdles e.g. for working with hundreds of thousands of files in single directories.

Availability: The package *boinc-server-autodock,* distributed with the experimental section of Debian Linux, provides scripts for an automated setup of complete desktop grid-ready drug screening project.

Acknowledgements: The authors thank Stefano Forli and his colleagues at The Scripps Research Institute, Gianfranco Costamagna and Antony Travis of Debian Med, Minke Informatics, and all our colleagues contributing to our BOINC docking projects.

♣ The work is supported by a DAAD fellowship to N.N. Nikitina and grant RFBR 13-07-00008, "Research and development of mathematical models and algorithms for creation of adaptive BOINC-based Desktop Grids".

A STRUCTURAL MECHANICS MODEL FOR ATOMIC FORCE MICROSCOPY-BASED INDENTATION TEST OF EPIDERMAL PLANT CELLS

Nikolaev S.V.[1]*, Golushko S.K.[2]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia*
e-mail: nikolaev@bionet.nsc.ru
**Corresponding author*

Key words: COMSOL, multiphysics, finite element method, FEM, mathematical modelling, structural mechanics, atomic force microscopy, AFM, plant cell wall

Motivation and Aim: Growth, development, and physiological reactions of plants are strongly based on the cell wall mechanics. One of the new experimental technique to determine mechanical properties of the plant cell wall is atomic force spectroscopy. But, in our point of view, there are some controversies in published results. To clarify mechanical behavior of the plant cell wall under experimental conditions we developed a structural mechanics model of the experiment to perform numerical simulations of this one.

Methods and Algorithms: We suppose that testing part of cell wall is its upper part which looks as a dome with rectangular base on epidermis. We considered two types of material for the dome: isotropic and orthotropic. To model cell internal medium we considered that the dome is under pressure from its inside surface. In some calculations we supposed the pressure is constant during indentation, while in other ones we considered the medium inside the dome is uncompressible. We formulated the all variants of experimental protocols using the Structural Mechanics Mode in the COMSOL 4.3b package. And we used the COMSOL facilities to perform calculations, and to analyze and to picture the results.

Results: Numerical simulations of the experimental testing of the plant cell wall mechanical behavior under nano- and micro-indentation demonstrated, that it is dependent of many parameters, for example material Young modulus, cell wall thickness, and its curvature, internal (turgor) pressure, and compressibility of cell internal medium.

Conclusion: From comparison of the computational experiments with published AFM-based testing of plant cells it is clear, that interpretation of the natural experimental results should be made with using numerical simulations. In particular, it is useful for correct estimation of cell wall material mechanical properties.

Acknowledgements: This work was supported by programs of the Presidium of the Russian Academy of Sciences nos. 6.6, SBRAS integration projects 47, 130. The authors are grateful to the Supercomputing center of the Novosibirsk state university for the opportunity to use COMSOL 4.3b.

A UNIVERSAL MODEL OF THE EPIDEMICS CAUSED BY SPECIAL PATHOGENS: A TOOL FOR EPIDEMIOLOGIST

Nizolenko L.Ph.*, Bachinsky A.G.
State Research Center of Virology and Biotechnology «Vector», Koltsovo, Russia
e-mail: nizolenko@vector.nsc.ru
**Corresponding author*

Key words: Epidemics, Mathematical model, Limited resources

Motivation and Aim: The outbreaks of infectious diseases caused by natural factors or bioterrorism acts are, quite real threats for the overall population. Correspondingly, it is necessary to elaborate mathematical models describing dynamics of the socially significant epidemics and epidemics caused by the most important special pathogens and the effects of antiepidemic efforts.

Methods and Algorithms: A universal deterministic model for predicting scenarios of epidemics under conditions of limited resources is developed. This model is presumably able to describe the dynamics of any epidemics of an acute disease caused by infection from a certain external source or via random contacts as the main transmission routes. Detailed description is done in [1].

Results: The model is equipped with a web interface and can provide comfortable job for an epidemiologist, who has no modeling experience. It allows:

- to select an infectious agent from the default list (influenza, smallpox, anthrax, plague, Ebola , Marburg, Lassa and Crimean-Congo hemorrhagic fevers) or to specify epidemic parameters for agents absent from the list of modeled agents;
- to get access the review of published data on epidemiology, pathogenesis, prevention, and treatment of diseases from the list of modeled ones;
- to edit parameters that specify the features of infection, the time of switching on antiepidemic activities and their intensity, the population size and resource availability for a region involved in simulation, and other;
- to compute the dynamics of epidemic; view/save the results;
- to calculate the losses caused by epidemic and supplies needed;
- to compute the dynamics of epidemic from a selected day;
- to compute four typical scenarios in development of an epidemic in a certain region;
- to simulate an accidental infection with a view to take into account potential interest to study variants with small number of infected persons.
- to optimize utilization of resources when responding to epidemics using genetic algorithm.

Conclusion: Along with a broad scope of simulated infections and a variety of factors control epidemics, advantage of the model is that it focuses on the user who is not a specialist in the field of information technology, being designed to solve practical problems in the field of protection of the population from epidemics. It can be used by individual epidemiologists as well as regional organizations, whose activities include participation in the eradication of epidemics.

Availability: http://vector-epimod.ru.

References:

1. A.G. Bachinsky, L.Ph. Nizolenko. (2013) A Universal Model for Predicting Dynamics of the Epidemics Caused by Special Pathogens, BioMed Research International, Article ID 467078, 7 pp, http://dx.doi.org/10.1155/2013/467078.

SIBERNETIC: NOVEL APPROACH TO REALISTIC MODELING OF INVERTEBRATES BIOMECHANICS

Palyanov A.*, Khayrulin S.
A.P. Ershov Institute of Informatics Systems SB RAS, Novosibirsk, Russia
e-mail: palyanov@iis.nsk.su
**Corresponding author*

Key words: Sibernetic, OpenWorm, virtual *C. elegans*, simulation, biophysics, biomechanics, incompressible smoothed particles hydrodynamics, OpenCL

Motivation and Aim: Computational biology is asserting itself as an important approach to understanding complex biological systems, with due attention paid to highly-detailed realistic biological neural networks simulations. They, in turn, require the organism's body for efficient work – including sensory system to provide realistic input and muscular system for producing body movement. Building a robotic system operating in real world, coupled with software simulation of nervous system, is quite challenging. Participating in international project called OpenWorm, aimed to build the first comprehensive computational model of the *C. elegans*, tiny 1 mm long roundworm, we developed special-purpose 3D physical engine oriented to realistic computational simulation of invertebrates biomechanics, particularly *C. elegans*.

Methods and Algorithms: Our approach is based on the modification of smoothed particle hydrodynamics (SPH) method which allows simulation of incompressible fluid dynamics. Incompressibility is enforced by using a prediction-correction scheme to determine the particle pressures (PCI SPH) [1]. We implemented this algorithm using modern programming technique called OpenCL, which provides an ability to perform high-performance parallel calculations at both GPUs and multi-core CPUs using the same source code. Next, we supplemented it with more accurate and reliable boundary-handling algorithm [2]. And finally, we developed and implemented a number of original features which are necessary and oriented to simulations in the field of invertebrates biomechanics. Among them are elastic matter which can contract in response to incoming signal – for simulation of muscle tissue, and liquid-impermeable 'membranes' preventing liquid particles from passing through their surface, which can be used to model a variety of objects from single living cell membrane to full body surface of invertebrates like mollusks or worms.

Results: We built functional prototype of biophysical engine described above, called Sibernetic, which is, to our knowledge, the first open source, parallel OpenCL/C++ PCISPH high-performance software of its kind [3]. As a proof-of-concept, a 3D model of the *C. elegans* body was created within Sibernetic, including a body shell filled with liquid maintaining its form due to hydrostatic pressure and a full set of 95 body wall muscles. We showed that this model is capable of producing *C. elegans*-like swimming locomotion in a liquid environment.

Availability: Sibernetic is available at GitHub repository (http://sibernetic.org)

Acknowledgements: The work was supported by Russian Foundation for Basic Research, grant № 14-07-31039.

References:

1. B. Solenthaler, R. Pajarola. (2009) Predictive-corrective incompressible SPH. ACM Trans. Graph. 28(3): 40:1–40:6.
2. M. Ihmsen, N. Akinci, M. Gissler, M. Teschner. (2010) Boundary Handling and Adaptive Time-stepping for PCISPH. Proc. VRIPHYS: 79-88.
3. M. Vella, A. Palyanov, P. Gleeson, S. Khayrulin. (2013) Integration of predictive-corrective incompressible SPH and Hodgkin-Huxley based models in the OpenWorm *in silico* model of C. elegans. BMC Neuroscience, 14(Suppl. 1): P209.

THE HIGH-DIMENSIONAL MODELS IN SOME TASKS OF BIOLOGY AND MEDICINE: PROBLEMS OF ANALYTICAL AND NUMERICAL STUDIES

Pertsev N.V.
Sobolev Institute of Mathematics, Omsk Branch, Omsk, Russia
e-mail: homlab@ya.ru

Key words: individual-based models, cellular automata, system of differential, integro-differential and stochastic difference equations of high dimension, population dynamics, epidemiology, method of monotone operators, M-matrices, Monte-Carlo method, parallel computing

Motivation and Aim: One approach to the study of the dynamics of the different populations is associated with the use of individual-based models and cellular automata models. Such models allow a detailed account of the features of individuals, mechanisms of their local interaction, the influence of environmental conditions and other factors. Models of this type cannot be analytical study and require massive computational experiments using high-performance computers. In some tasks more promising is the use of models in the form high-dimensional system of differential, integro-differential and stochastic difference equations. In the framework of these models individuals are indistinguishable by some signs within the examined group. The number of such groups can be quite large, and the number of individual in groups may range from a few to hundreds, thousands, or more individuals. In the presence of small groups is advisable to use stochastic models with integer variables. The lecture will be presented examples of the following high-dimensional models: 1) deterministic model of the dynamics of a population that is developing in the conditions of influence of harmful substances; 2) deterministic model of the spread of tuberculosis among the population of several regions; 3) stochastic model of HIV infection on the basis of levels of social exclusion of individuals.

Methods and Algorithms: For studying solutions of high-dimensional models we use methods of the theory of monotone operators, the properties of nonsingular M-matrices, numerical algorithms of the Monte-Carlo technique and parallel computing.

Results: The lecture contains typical solutions of the considered models, including solutions interpreted as a degenerate populations, and solutions associated with the reduction of the incidence of tuberculosis and HIV infection.

Conclusion: The examples point to the efficiency of application of high-dimensional models in combination with modern apparatus of analytical and numerical methods of research.

References:

1. N. Pertsev, G. Tsaregorodtseva. (2011) Modeling population dynamics under the influence of harmful substances on the individual reproduction process, Automation and Remote Control, **72:1:** 129-140.
2. N. Pertsev, V. Leonenko. (2009) Stochastic individual-based model of spread of tuberculosis, RJNAMM, **24:4:** 341-360.
3. N. Pertsev, V. Leonenko. (2012) Discrete stochastic model of HIV infection spread within a heterogeneous population, RJNAMM, **27:5:** 459-477.

SOLUTION OF INVERSE IMMUNOAGGLUTINATION KINETICS PROBLEM FOR PATCHY PARTICLES WITH A SMALL NUMBER OF BINDING SITES

Polshchitsin A.A.*, Nekrasov V.M., Chernyshev A.V., Maltsev V.P.
Institute of Chemical Kinetics and Combustion SB RAS, Novosibirsk, Russia;
Novosibirsk State University, Novosibirsk, Russia
e-mail: polshchitsin@gmail.com
**Corresponding author*

Key words: immunoassay; aggregation kinetics; biospecific binding; immunoagglutination; discrete receptors; diffusion rate constant; Monte-Carlo simulation

Motivation and aim: Turbidimetric immunoagglutination method of the quantitative detection of specific proteins is widely used in clinical laboratories because of its cheapness, simplicity and affordability. However, this method has a number of limitations in accuracy of the results obtained due to the lack of detailed description of the processes taking place during the agglutination, caused by the complexity of the mathematical model of the biospecific aggregation kinetics and the difficulty of theoretical calculations of optical density change. This work is devoted to the construction of the model to describe in details the process of agglutination of polymer microspheres with the discrete number of binding sites.

Methods and algorithms: In this work we used Monte Carlo algorithm of numerical computation of stochastic evolution in time of paired chemical reactions for the solution of direct kinetic problem of patchy particle agglutination with few binding sites. We assume the initial distribution of particles in occupied binding sites being described with binomial distribution. Autocorrelated noise of Monte Carlo simulations was reduced by means of averaging over several realizations with fixed parameters. Optical features of particle aggregates were calculated using the superposition T-matrix method.

Results: We have solved the inverse problem for the example set for latex turbidimetric determination of the concentration of C-reactive protein in human plasma. Excellent agreement between theoretical and experimental curves in a wide range of experimental conditions, achieved through a global optimization method DiRect, has shown. As a result of the fitting the physical parameters of the system are determined. They coincide with literature data and numerical estimates within the errors of calculations.

Conclusion: The solution of the inverse problem using constructed algorithm of immunoagglutination kinetics calculations and superposition T-matrix method allows one to determine physical parameters of the examined system with high accuracy. For the patchy particles with few binding sites the extreme necessity of accounting the discreteness of reactive spots has been shown.

LOGICAL-AND-PROBABILITY SIMULATION MODEL OF DATE ANALYSES

Polyakova G.L.
Novosibirsk State University, Novosibirsk, Russia
e-mail: Polyakova@math.nsc.ru; Karusel2007@ya.ru

Key words: simulation, time series, event and decision tree, logical rules

Motivation and Aim: Logical-and-probability simulation model provides practical understanding of the methods of classification, prediction, reduction and data mining.

Methods and Algorithms: The methods of multimodal adaptive optimization for searching the optimal decision tree is proposed. New method of construction of logic-and-probabilistic models for multidimensional heterogeneous time series is proposed. The method is based on adaptive algorithm of multimodal optimization. The approach is based on searching partitions of informative variables space that in the best way separate observations with different levels of outcomes. Statistical validity associated with revealed regularities is estimated with the help of permutation test repeating search of optimal partition for each permuted dataset. Effectiveness of proposed method was investigated.

Results: The methods were applied to solving several real tasks in ecology and bioinformatics. A decision tree was built by partitioning the space variables with the help of probabilistic algorithm that is moved from the current solution to the next solution that is sampled uniformly from the set of all better feasible solutions. Obtained solution was optimum by criterion of minimizing the risk of erroneous classification. They are especially useful in areas, where few data can be collected.

Conclusion: The algorithm was used for the following applications: analysis of environmental factors influencing the incidence of tick-borne encephalitis in endemic regions of Russia. Suggested methods have some positive features. For example, obtained results are similar to natural language and reflect cause-effect relations of objects under investigation.

References:

1. G.S. Lbov, M.K. Gerasimov, G.L. Polyakova (2010) A Method of Interval Prediction based on Logical Regularities, Proc. of the 3rd IASTED Intern. Conf. on Automation, Control, and Information Technology (ACIT 2010). 15-18 June 2010, Novosibirsk, Russia. 2: 245-249.
2. G.S. Lbov, G.L. Polyakova (2011) Forecasting Method Based on Logical Regularities, *Proc. of the 11th International Conference on Pattern Recognition and Information Processing (PRIP'2011). May 18-20, 2011, Minsk, Belarus. Minsk: BSU*: 317-319.
3. V.B. Berikov, G.S. Lbov, G.L. Polyakova et al. (2011) Analysis of factors influencing the incidence of tick-borne encephalitis, using logical-and-probabilistic and correlation-regression models, *Epidemiology and Immunization,* 6(61): 25-33.
4. G.S. Lbov, G.L. Polyakova, V.N. Bakhvalova, O.V. Morozova (2010) Study of the influence of natural factors on incidence of tick-borne encephalitis, *Bulletin of the NSU. Series: biology, clinical medicine,* 8(2): 31-37.

LOGICAL-AND-PROBABILITY SIMULATION MODEL OF INFLUENCE OF THE CLIMATIC FACTORS IN TICK-BORN ENCEPHALITIS DISEASE'S INDICES

Polyakova G.L.
Novosibirsk State University, Novosibirsk, Russia
e-mail: Polyakova@math.nsc.ru; Karusel2007@ya.ru

Key words: event and decision tree, logical rules, tick-born encephalitis

Motivation and Aim: Tick-borne encephalitis (TBE) is considered an international health issue, as the number of risk areas and reported cases across Europe, Russia, and parts of Asia continues to increase. The risk for contracting TBE depends strongly on the density of the infected questing ticks and many studies have investigated tick population dynamics and the parameters affecting them. To understand the fluctuation in incidence rates during the period of observation a complex interrelation of several factors has to be considered, such as astrophysical and climatic factors. It can be assumed that changes of leisure activities in nature, increasing/decreasing mobility to risk areas, changes in wildlife hosts/tick populations may have influenced the quantity and quality of epidemiological data. In this paper a logical-and-probabilistic simulation model for predicting epidemiological data is presented.

Methods and Algorithms: Joint analysis of multidimensional time series of annual TBE prevalence among Novosibirsk, Irkutsk and Gorno-Altaysk populations, monthly air temperatures, relative humidity, rainfalls, annual solar activity expressed as Wolf's numbers and vaccination rate was performed using logical-and-probabilistic models. In order to determine the general relevant factors was performed joint information processing on all the regions. Analysis of the influence of natural factors of the current year (last 12 months) for the value of Y, and the Y forecast in the current year were carried. When processing the raw data for the variable Y was selected three intervals, there are three images. The first image was determined by the size of the TBE incidence up to 10 (low incidence) and the second from 10 to 20 (average), the third is greater than 20 (high). A decision tree was built by partitioning the space variables with the help of probabilistic algorithm that is moved from the current solution to the next solution that is sampled uniformly from the set of all better feasible solutions. Obtained solution was optimum by criterion of minimizing the risk of erroneous classification.

Results: By means of decision tree were revealed common significant factors for three studied TBE natural foci. As a result, more 40 logical regularities were obtained [1, 2], for example, relative air humidity of November of preceding year, relative humidity of April and June as well as temperature of June of current year. Unidirectional positive correlation between solar activity expressed in Wolf's numbers and TBE rate was shown for Novosibirsk and Irkutsk but for Gorno-Altaysk such correlation after 2001 was absent.

Conclusion: The analysis of environmental factors influencing the incidence of tick-borne encephalitis in endemic regions of Russia confirmed the early results for Europe. The correlation between the decrease of questing ticks in the summer and the combination of air temperatures and humidity in the form of saturation deficit. However, up until now there has been no single method that can tackle the problem of influence of climatic factors on dynamics ixodes ticks in all its generality and completeness. The field is, thus, still open to research.

References:

1. V.B. Berikov, G.S. Lbov, G.L. Polyakova et al. (2011) Analysis of factors influencing the incidence of tick-borne encephalitis, using logical-and-probabilistic and correlation-regression models, *Epidemiology and Immunization,* **6(61):** 25-33.
2. G.S. Lbov, G.L. Polyakova, V.N. Bakhvalova, O.V. Morozova (2010) Study of the influence of natural factors on incidence of tick-borne encephalitis, *Bulletin of the NSU. Series: biology, clinical medicine,* **8(2):** 31-37.

SHARED BIOINFORMATICS DATABASE WITHIN UNIPRO UGENE

Protsyuk I.V.[1, 2]*, Grekhov G.A.[2], Tiunov A.V.[2], Fursov M.Y.[2]
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Center of Information Technologies «UniPro», Novosibirsk, Russia*
e-mail: iprotsyuk@unipro.ru
**Corresponding author*

Key words: Bioinformatics, Custom Shared Databases, Big Data, Open Source, Cross Platform, Data Management

Motivation and Aim: Unipro UGENE [1] is an open-source bioinformatics toolkit that integrates popular tools along with original instruments for molecular biologists within a unified user interface. Nowadays most bioinformatics desktop applications, including UGENE, make use only of local user files when processing different types of data. Such an approach causes inconvenience to scientists working cooperatively and relying on the same data. The most obvious issues are the need to make multiple copies of certain resources for every workplace and do synchronization in further. There are tools that provide the needed capabilities but they are proprietary and quite expensive.

Methods and Algorithms: It was decided to implement the client-server architecture where the server is represented by a remote database server and the client is an instance of the UGENE tool run on a local host. Thereby certain data models were developed for all the types of biological data supported by the UGENE platform. One of the most important goals was to minimize the count of queries to the server. Major efforts were made to avoid downloading of the entire data, requested by a user as well. Loading on demand was realized instead.

Results: Initial tests showed that the system is capable of storing tens of thousands of objects (annotated sequences, multiple alignments, etc.) with shared read/write access to them. Large objects are supported as well. For example, the assembly of Denisovan genome (~10 GB in BAM format) was successfully imported to the database and clients could navigate through its short reads.

Conclusion: Thus, UGENE may provide a unified access to shared data for users located, for example, in the same lab or an institution. This work itself presents a basis for further development of a distributed computational system. Namely, workflow manager, a part of the UGENE platform, coupled with a shared database may operate remotely and do jobs for clients aiming to process the shared data. Therefore, this opens an opportunity for using easy-to-install computational clusters requiring the UGENE suite only.

Availability: UGENE binaries are freely available for MS Windows, Linux and Mac OS X at {{http://ugene.unipro.ru/download.html}}. UGENE code is licensed under the GPLv2.

References:

1. Konstantin Okonechnikov et al. (2012), Unipro UGENE: a unified bioinformatics toolkit, Bioinformatics, 28: 1166-1167.

MODELLING OF THE PROBLEM OF MULTIPLE ALIGNMENT OF THE NUCLEOTIDE SEQUENCES AND DENDROGRAM CONSTRUCTION

Pyrkova A.Yu.*, Ivashchenko A.T., Berillo O.A.
Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: Anna.Pyrkova@kaznu.kz
**Corresponding author*

Key words: algorithm, alignment, dendrogram, Java MPI, mathematical model, miRNA

Motivation and Aim: The detection of the quantitative characteristic of similarity of nucleotide and aminoacid sequences is often problem in genomics and proteomics [1]. The difficulties of the solution of this problem is a great number of RNA sequences and their length. One of the disadvantages of the dynamic programming algorithms is their enough low speed of alignment. The alignment of sequences more than one million symbols or performance of hundreds thousands comparisons of sequences consisted of some thousands symbols is impossible to realize for acceptable time even on the modern computers. Aim of this research is optimize decision of the problem of multiple alignment of the nucleotide sequences and dendrogram construction using the computational cluster.

Methods and Algorithms: For achievement more exact result of multiple alignment and optimization of the time demanded for data processing at multiple alignment, the set of breaks of sequences on the M independent groups processed by M parallel processes. Each of them carries out the alignment of its group independently of others.

Results: The mathematical model of multiple alignment of the nucleotide sequences was constructed. The algorithm of multiple alignment constructed on the basis of the analyzed Needleman-Wunsch algorithm which was modified for processing of files with big data via parallelization using MPJ. The parallelized algorithm of dendrogram construction was realized on the cluster computing platform. The results of work of the program were tested on the nucleotide sequences of microRNAs (miRNAs). Several hundreds and thousands the short miRNA nucleotide sequences with length from 17 to 27 nucleotides was compared. They consist of only four different symbols. Their specific length and content represent a complex problem. The adequate solution of this issue is necessary for identification miRNA binding sites with mRNAs. In this case the received characteristics helped to establish belonging miRNA to one or different families. Interpretation of biological function miRNAs can be depended on the solution of this question. The membership of some miRNAs of Arabidopsis to one family was established by created program. For example, miR156 and miR157 are members of one family as well as ath-miR171 and ath-miR171a belong to another family. The developed program will be applied to distribution 2500 human miRNAs to different families.

Conclusion: Different types the programs can be demanded in process of a sequencing of genomes of animals and plants. The developed algorithm used for analysis of miRNAs of different organisms.

Availability: The program works on the computing cluster of al-Farabi Kazakh National University (http://ursa.kaznu.kz/).

References:

1. N.C. Jones, P.A. Pevzner. (2004) An Introduction to Bioinformatics Algorithms. (Massachusetts Institute of Technology Press, Massachusetts).

PARALLELIZATION OF ALGORITHM OF PREDICTION OF miRNA BINDING SITES IN mRNA ON THE CLUSTER COMPUTING PLATFORM

Pyrkova A.Yu.*, Ivashchenko A.T., Berillo O.A.
Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: Anna.Pyrkova@kaznu.kz
**Corresponding author*

Key words: parallelized algorithm, cluster computing platform, Java MPI, miRNA, mRNA

Motivation and Aim: The problem of a prediction of binding sites of microRNAs (miRNAs) with messenger RNAs (mRNAs) had increased after discovery of an important role of miRNAs in regulation of gene expression. There are data about the value of free energy of hydrogen bond between nucleotides in water solution [1]. However, there are a wide range of free energy value of this bond and it is difficult to choose a correct data. It is important to know the relative relations of free energy of hydrogen bond between nucleotides as they are necessary at formation of secondary and tertiary structures of RNAs. Some programs which predict miRNA binding sites were created, however, many of them had unreasonable limitations for search of binding sites. It was established that binding sites are localized only in 3'UTRs. Other programs were based on identification of binding sites with the obligatory requirement to have complementary interactions. Many such programs predict a large number of false positive sites and they are not allow to reveal the binding sites located in 5'UTRs and CDSs. Aim of this research is creation of the program which has not aforementioned disadvantages and with high reliability revealed binding sites of miRNAs with mRNAs.

Methods and Algorithms: Scanning genes is a process of consecutive comparison of nucleotide sequence of mRNA with miRNA with possibility of adding one gap in miRNA sequence in positions with the 3-rd on n-2-th, where by n – nucleotide number (length) of miRNA. The binding sites are selected according to the value of free energy of compared sequences. It is considered the best that option which is closer (in a percentage ratio) of free energy for coincidence of miRNA with binding site on the basis of complementarity.

Results: The developed algorithm scans of mRNA with miRNA with one possible gap in miRNA sequence, calculates a maximum of free energy and analyzes a coincidence of miRNA and binding site of mRNA on the basis of complementarity properties. The algorithm is base of MirTarget program. The program determines a free energy of miRNA hybridization with mRNA and the schemes of their interactions. It calculates of the nucleotide bounds of binding sites, level of reliability and the mRNA parts, where the sites are located since the first nucleotide. The algorithm can be parallelized on the computational cluster with use of MPJ tools.

Conclusion: The MirTarget program has advantages which are not present in known programs predicting of miRNA binding sites with mRNAs.

Availability: The program works on the computing cluster of al-Farabi Kazakh National University (http://ursa.kaznu.kz/).

References:

K.M. Guckian, B.A. Schweitzer, R.F. Ren, C.J. Sheils, P.L. Paris. (1996) Experimental measurement of aromatic stacking in the contex to duplex DNA, Vol. 118: 8182–83. (J.Am.Chem.Soc.).

DOTOLOG – DOT PLOT ANALYZATION TOOL

Romanov D.E.*
A.B. Kogan Research Institute for Neurocybernetics Southern Federal University, Rostov-on-Don, Russia
e-mail: rdme@ya.ru
**Corresponding author*

Key words: dot-plot, homology

Motivation and Aim: Dot plots are one of the oldest ways of comparing two sequences. These technique was introduced by Gibbs and McIntyre in 1970 [1]. In bioinformatics a dot plot is a graphical method that allows the comparison of two biological sequences and identify regions of close similarity between them. It is a kind of recurrence plot.

The problem is that the modern biology still lack of strict definition of the term "homology". On the other side, the expert could rather fast make up a conclusion about the degree of the homology between two sequences, as far as a proper dot-plot image is shown.

Despite this technique of analysis does need expert labor opposite to automatic sequence comparison with help of various alignment tools, like blast [2] or glam2 [3], it may be much faster and accurate. Also this method allows direct observation of the basic types of sequence rearrangements, like insertions, deletions, translocations or duplications.

Methods and Algorithms: As dot plot analysis requires scientist's labor, the special attention was paid to ergonomics and ease of handling. The program requires Java Runtime Environment.

Interface implements "as Google maps" gestures (mouse dragging with left button and scaling with mouse wheel) with some additional features. Color coding is widely used, e.g. UTR features are highlighted in dark red, CDS features - in dark blue and words are highlighted in green.

The main innovation is the ability to select some motif. Selected motifs are highlighted in dark cyan. Also the detailed description of the motif immediately appears in text area at the bottom of the screen. The output format is rather simple and represents the variation of CSV format.

With increasing of the scale, more detailed plot is available. It is also possible to highlight all matches in selected area.

For convenience, the visible part of main screen can be imported to raster format (PNG) for later manipulations.

Conclusion: This tool was developed as part of large sequence analyzation pipeline.

Even for relatively large sequences (approximately 30000bp), visual analysis could be done for only 5-10 minutes.

Availability: The programm is still in active development, and the latest version is available at http://biosfedu.no-ip.org/progs/.

References:

1. Gibbs, A.J. The diagram, a method for comparing sequences. Its use with amino acid and nucleotide sequences. [Text] / A. J. Gibbs, G. A. McIntyre // Eur J Biochem. 1970. Sep. Vol. 16, no. 1. P. 1-11.
2. Basic local alignment search tool. [Text] / S. F. Altschul, W. Gish, W. Miller [et al.] // J Mol Biol. 1990. Oct. Vol. 215, no. 3. P. 403-410. URL: http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
3. Discovering sequence motifs with arbitrary insertions and deletions. [Text] / Martin C. Frith, Neil F W. Saunders, Bostjan Kobe, Timothy L. Bailey // PLoS Comput Biol. 2008. Apr. Vol. 4, no. 4. P. e1000071. URL: http://dx.doi.org/10.1371/journal. pcbi.1000071.

DATABASE OF FRAME MODELS OF GENETIC REGULATION OF THE METABOLIC PROCESSES ASSOCIATED WITH DISEASES

Saik O.V.[1]*, Demenkov P.S.[1], Ivanisenko V.A.[1]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
e-mail: saik@bionet.nsc.ru
**Corresponding author*

Key words: database, metabolites, genes, enzymes, transcription factors, miRNAs, SNPs, diseases, genetic regulation, frame associative models, metabolic processes

Motivation and Aim: Investigation of molecular genetic mechanisms of diseases gives an opportunity to identify new potential drug targets and diagnostic markers. Reconstruction of frame associative models of genetic regulation of metabolic processes associated with diseases can help to address this problem. Given the abundance of information on metabolism, its regulation and associations with diseases in different data sources the use of automatic data extraction and analysis is very relevant. The aim of this work was to create a database of frame models of genetic regulation of metabolic processes associated with diseases using text- and data-mining.

Methods and Algorithms: Information from the databases HMDB, KEGG, ChEBI, MetaCyc, UniProt, Brenda, TRRD, dbSNP, SNPedia, miRTarBase, miR2Disease, HMDD was used. Text-mining was performed by ANDSystem (Demenkov *et al*., 2012).

Results: Promedia database containing more than 150 million of frame models of genetic regulation of metabolic processes associated with diseases was developed. A frame model is a sequence of metabolites, enzymes (and their genes), transcription factors (and their genes), microRNAs, SNPs related to each other and diseases. Example of one type of frame models with SNP in the transcription factor (TF) gene:

SNP → TF gene → TF → enzyme gene → enzyme → metabolite
| | | | | |
disease disease disease disease disease disease

The developed database provides information on 7983 metabolites, 512 transcription factors, 7042 enzymes, 3274 microRNAs, 11165 SNPs, 3945 diseases their relationships and properties (Saik *et al*., 2011).

Conclusion: Promedia database containing more than 150 million of frame models of genetic regulation of metabolic processes associated with diseases providing an opportunity to predict new potential drug targets and diagnostic markers was developed.

Availability: Promedia is available at http://www-bionet.sscc.ru/promedia/

Acknowledgements: this work was partially supported by the projects VI.61.1.2., НШ-5278.2012.4 (Программа РАН «Молекулярная и клеточная биология»).

References:

1. Demenkov P.S. *et al*. (2012) ANDVisio: A new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem, *In silico biology*, **11**(3): 149-161.
2. Saik O.V. *et al*. (2011) Database of Chemical Substances the Potential Biomarkers of Diseases with the Meaning in Noninvasive Diagnostics, *Mathematical Biology and Bioinformatics*, **6.2**(2011): 250-263.

MODELLING OF SOFT TISSUES DEFORMATION. ALTERNATIVE APPROACHES

Salamatova V.Yu.*
Moscow Institute of Physics and Technology, Moscow, Russia
e-mail: salamatova@inbox.ru
**Corresponding author*

Key words: soft tissue, mathematical modelling

Motivation and Aim: Real-time simulations of medical procedures caused development of alternative methods for soft tissues deformation. These methods should be effective in terms of computational speed, accuracy and real-time performance.

Methods and Algorithms: There are a bunch of different methods and approaches for modelling soft-tissue deformation, and they can be divided into two groups: mesh-based and meshless methods. Frame-based elastic models and Mass-Spring models will be discussed in detail.

Results: Two methods of alternative approaches to modelling soft tissues deformation can be discussed.

Conclusion: The direct application of the classical methods such as finite element method can be time-consuming and difficult. In addition, due to lack of accurate data on tissue mechanical behavior, using complex formulations that sacrifice computational velocity can lead to simulations that are not more accurate than those achieved with simpler models. Thus, the alternative methods for modelling of soft tissues deformation need to be further developed.

Availability: It is a review; there is no need of any software.

CONTINUATION OF THE ACOUSTIC FIELD IN TOMOGRAPHY

Shishlenin M.A.
Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;
Novosibirsk State University, Novosibirsk, Russia
e-mail: mshishlenin@ngs.ru

Key words: inverse and ill-posed problem, regularization methods, numerical methods, continuation problem

Motivation and Aim: We consider the continuation problem of the acoustic field inside the domain. Such problems can be considered in medicine tomography, when continuation of physical fields allows to detect inclusions in the investigated domain. The continuation problems are related to inverse problems of mathematical physics [1].

Methods and Algorithms: We reduce the ill-posed problem to the operator equation [2,3]. For numerical solution of the continuation problem we apply singular value decomposition method and gradient method.

Results: Theory and numerical methods are developed for the continuation problem of the acoustic field. As the result, the formulae to calculate the singular values of the continuation problem operator have been obtained.

Acknowledgements: The work was supported by integration projects 14 of SB RAS and RFBR grant 12-01-00773.

References:

1. S.I. Kabanikhin. Inverse and Ill-Posed Problems. Theory and Applications. De Gruyter, Germany, (2011).
2. S.I. Kabanikhin et al. (2013) Inverse Problems for the Ground Penetrating Radar, *Journal of Inverse and Ill-Posed Problems*, **21** (6).
3. S.I. Kabanikhin et al. (2013) Regularization of the continuation problem for elliptic equations, *Journal of Inverse and Ill-Posed Problems*, **21** (6).

MICROSATELLITE VARIATION TO COMPARE MIGRATION SCENARIOS AND DEMOGRAPHIC PROCESSES IN POPULATIONS CHUM SALMON NORTHERN RANGE

Shitova M.V.[1]*, Khokhlov Y.N.[2], Nikiforov A.I.[3], Sherbakov D.Y.[4], Zhivotovsky L.A.[1]
[1]Vavilov`s Institute of General Genetics RAS, Moscow, Russia;
[2]FSUE «ChukotTINRO», Anadyr, Russia;
[3]Moscow State Institute of International Relations (University) MFA RF, Moscow, Russia;
[4]Limnological Institute SB RAS, Irkutsk, Russia
e-mail: shitova-m@rambler.ru
**Corresponding author*

Key words: microsatellite variation, population genetics, coalescent analysis

Motivation and Aim: In publications devoted to population-genetic studies, there are sections on the testing of a hypothesis about the ways of formation of the modern genetic diversity. There are software packages on the calculation of population and demographic parameters (both contemporary and historical) - the historical dynamics of the effective population size, historical dynamics of migration flows between populations, the time of separation of populations, etc. In this connection it is necessary to accumulate data model and to compare them with the actual population-genetic parameters.

Methods and Algorithms: We used the principles coalescent analysis based on the Monte Carlo Markov chains (MCMC), coalescent analysis program - Migrate-n. We used variability of 10 microsatellite loci in three populations of chum salmon northern part area.

Results and conclusion: (first) In Chukotka chum salmon population is tested growth of effective population size, which started in the recent past - this fact will be checked for stability and repeatability of results from run to run; (second) We detected sporadic bursts of effective population size and bursts of migration between pairs of populations - data bursts is not very stable and periodically displaced in time - these bursts will be refined and tested for reproducibility. Reported features 1 and 2 may be associated with paleoclimatic history of the region (northern Pacific); (third) We defined relative mutation of rates in each of the 10 microsatellite loci (minimum for the locus Oki1-2 - 0.28777, maximum for the locus One103 - 2.23022). Variation in mutation rate was fairly broad in terms of absolute values (relative values must be multiplied by 6,2 × 10-4), mutation rates differ by an order; (fourth) were tested: (1) a complete model of migration - migration occur between all populations in all directions, (2) model, where migration is totally prohibited between Alaska and riv.Penzhina, (3) migration is totally prohibited between Chukotka and Alaska, (4) the migration is completely prohibited between Chukotka and riv.Penzhina. Testing has shown that the most likely to be full migration model (№ 1).

ON MATRYOSHKAS AND BIOMEDICAL PROTOCOLS

Soldatova L.N.[1, 2]*, Nadis D.[2], King R.D.[2], Basu P.S.[2], Haddi E.[2], Baumlé V.[2], Saunders N.J.[1], Marwan W.[2], Rudkin B.B.[2]
[1]Brunel University, Uxbridge, UK;
[2]The University of Manchester, Manchester, UK
e-mail: larisa.soldatova@brunel.ac.uk
**Corresponding author*

Key words: biomedical protocols, Petri net

An ontology EXACT (EXperimental ACTions) defines the minimum information required for the recording of biomedical protocols. It is designed as a reference model for the representation of experimental actions and their essential properties. To construct EXACT we manually inspected hundreds of published and commercial biomedical protocols from several areas of biomedicine, including neurology, epigenetics, metabolomics, and stem cell biology. EXACT has been used for the translation of biomedical protocols from natural language to a representation in a semantically defined machine processable format. We also suggested an integration of the EXACT formalism with the Petri net formalism to enable the representation of experimental workflows. While EXACT captured the semantic meaning of experimental actions, Petri nets modelled sequences of such actions, including branching and loops. Petri nets thus provide a machine-readable interactive visualisation of experimental workflows, enabling one to collapse or expand specified parts of a workflow. It is possible to combine several experimental procedures within one Petri net if required, and to support a nested structure of an experimental workflow – we named it a "matryoshka" protocol.

METHODS OF THE INTELLECTUAL ANALYSIS OF DATA IN NATURAL SCIENCES AND ALTERNATIVE ANALYSIS

Sukhotin A.M.*
Physical-Technical Institute NRTPU, Tomsk, Russia
e-mail: asukhotin@yandex.ru

Key words: alternative series, C–pair of natural variables, e-divergent sequence, w-convergent sequence, Euclidian axiom 8, infinitely large number, a quantity of all prime numbers

Motivation and Aim. The main Aim of this study to obtain *alternative methodology and new result for an* analysis of data in natural sciences.

At the first we proved ***Theorem 1*****:** *Any permutation of alternative series addends does not change its convergence*.

Then we proved that a mapping $\varphi: N \to N$, $k = \varphi(n) \triangleq 2n$ *is impracticable* on all set N, in particular, we proved following ***Theorem 2.*** $\varphi(N) = N \Rightarrow \lim_{n\in N}(\varphi(n):n) = 1$.

The written down statements do not depend on Theorem 1 and its consequences. But the alternative methodology unites proofs of the first and the second.

Methods and Algorithms: New concepts and method мотивированной hypotheses form new alternative methodology. There are examples in ouradstract.

Definition 1. The pair (m, k) of variables $m\in A$, $k\in B$ is said to be *C–pair*, if there $\exists C\in N$ that every neighboring in $E \triangleq A \cup B \subseteq N$ elements $m\in A$, $k\in B$ hold the inequality $|m\text{-}k|<C$.

Definition 2. The number sequence (***a***) is named *e-divergent* one (*e-**DS***) if there are such two infinite subsequences $\xi_1, \xi_2 \subset N$, $\xi_1 \cap \xi_2 = \varnothing$, that there is fairly a following condition: $\exists(\delta>0, n^*\in N)$: $\forall(m, k)\in(\xi_1, \xi_2)$ $m, k > n^*$ $|a_m - a_k| \geq \delta$.

Definition 3**.** The number sequence (***a***) is said to be *w-convergent* **(*w*-CS)** if it satisfies following condition $(\forall\varepsilon>0\ \exists n(\varepsilon)\in N)$: $\forall n \geq n(\varepsilon)|a_m - a_k| < \varepsilon$, or $\lim(a_m - a_k) = 0$.

Remaining within the framework of sets naive theory by P. Halmosh we proved Euclidian Axiom 8 “The Whole is more its own Part” in the form of

Theorem 3**:** $B\subset A \Rightarrow (\forall\varphi: A \to B\ \exists(a, q)$, $a\in A$ and $q\in A$: $a\neq q$ & $\varphi(a) = \varphi(q)$.

Theorem 3 has *canonical brief form*: $B\subset A \Rightarrow \neg(A\sim B)$.

Now we formulate by [1] some *main Results:* ***Theorem 4.*** *There exist in the theory of number sequences some Cauchy sequence that unlimited final number*.

Now we enter a concept of infinitely large nuAmber correctly and show some their application.

The limit value of unlimited by whatever finite number Cauchy sequence (***a***) *is named an* ***infinitely large number*** (***ILN***), *defined by this sequence* (***a***).

Theorem 5. *There exists some* ***ILN*** Ω_π, *which defines a quantity of all prime numbers.*

Theorem 6. The harmonic series converges to some ***ILN*** Ω_e.

Theorem 7. An assumption, that the mapping $f(n) \triangleq n + 1$ defines the bijection $f: N \to N\backslash\{1\}$, implies to the contradiction. Hence, we have proved Euclidian Axiom 8 at the second time.

References:

1. Сухотин А.М. Начало высшей математики. -учеб. пособие. –3-е изд.–Томск: Изд-во ТПУ, 2008.– 164 с.

INDIVIDUAL MODELLING OF HEMODYNAMIC PROCESSES IN CARDIOVASCULAR SYSTEM BASED ON PERIPHERAL ARTERIAL PULSATION

Sviridova N.V.*, Vlasenko V.D.
Computing Center Far-Eastern Branch Russian Academy of Science, Khabarovsk, Russia
e-mail: nina.v.svr@gmail.com
**Corresponding author*

Key words: hemodynamics, cardiovascular system, peripheral arterial pulsation

Motivation and Aim: The blood circulation and performance of the cardiovascular system (CVS) has been an object of many recent studies, due to promising applications for human health care and monitoring. Most of investigations are based either on modeling or on biological signals analysis. Both approaches have various advantages, such as repeatability, controllability of numerical experiments and many others for mathematical models; and possibility to study individual features of CVS in each particular case or conduct real time health monitoring in clinical settings, etc. for signal analysis. Developing of numerous commercially available medical and sports devices for CVS health control bring up new issue in CVS monitoring development. As long as monitoring conducted only on the base of signal analysis only, prediction of CVS performance or reconstruction of parameters other than measured ones become a challenge. At the same time modelling allows forecasting and reconstruction but with a certain generalization, that isn't suitable for each individual case.The ultimate goal of this study was to develop method utilizing advantages of both mathematical modelling and signal analysis for signal based individual modelling.

Methods and Algorithms: The method was developed on the base of integration peripheral arterial pulsation (PAP) into well-known system of equation of hemodynamics in blood vessels [1]. This integration allowed us to conduct individual modelling by utilizing subject's PAP. PAP can be measured noninvasively by a simple and low-cost optical technique that can detect blood volume changes in the microvascular bed of tissue.

Results and Conclusion: Developed PAP-integrated individual model was applied for calculation of blood flow parameters based on PAP signal obtained from healthy 25 years old subject. Results showed certain degree of reliability and were in good agreement with results of studies utilizing either mathematical models or signal analysis only. Therefore developed individual model can be applied for CVS health monitoring.

References:

1. E.V. Astrakhantseva, et al. (2005) Mathematical modelling of hemodynamics of large blood vessels, *Mathematical modelling*, **17:** 61–80.
2. J. Allen. (2007) Photoplethysmography and its application in clinical physiological measurement, *Physiological measurement*, **28:** 1-39.

NETINFERENCE: THE COMPUTER TOOLS FOR ANALYSIS AND VISUALIZATION OF NETWORKS STRUCTURE, DYNAMICS AND EVOLUTION

Titov I.I.[1, 2]*, Blinov A.A.[2], Rudnichenko K.A.[2], Krutov P.V.[2], Kazantsev A.L.[2], Kulikov A.I.[2, 3]

[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia*
e-mail: titov@bionet.nsc.ru
**Corresponding author*

Key words: gene network, co-authorship network, term network, network structure, network dynamics, synchronous Boolean model, computer analysis

Motivation and Aim: Progress in modern science often depends upon exploratory and confirmatory data analysis, inference and modeling. These activities could eventually permit the prediction and control of complex systems. The computer tools which can help to visualize and to analyze the networks will play an essential role in this progression.

Methods and Algorithms: The gene networks were taken from open sources. The PubMed database accumulates information on the scientific papers in the medicine and biology and allowed us to build the co-authorship and word networks. The computational analysis of statistical and dynamical properties of networks was performed by the program package NetInference which is implemented in C++, C#, SQL and Java. The package consists of three programs. The first tool analyzes global and local network architecture by calculating degree distribution, correlation coefficients and other network characteristics, by motif search, etc. The second one simulates a gene network dynamics within synchronous Boolean model. The third program reveals clusters in a network and traces their evolution.

Results: We start this presentation from a gene network analysis. The key regulators are revealed by exploring the dynamical properties and by stability analysis. The network architecture is characterized by calculating the statistical characteristics and by revealing the network "computational core". The latter method is based on the network decomposition into the weakly interacting modules and by the network reduction.

Next we apply the package to analyze the evolution of the co-authorship network of biomedical research in Novosibirsk Scientific Center. We observe both stationary and non-stationary evolution of scientific groups inside the institutes. We build the inter-institute interaction matrix which allows us to characterize the intra-Novosibirsk and the international cooperation.

Finally we demonstrate how our computer tools could be used to map the scientific landscapes using the term networks. A typical scientific field experiences the times of "rise" and "fall" similar to the market bubbles and can be characterized by its "age". Specifically we explore the hot topics of nanotechnology and miRNA research.

Conclusion: We present a computer package for analyzing the structure-functional organization and evolution of biological, social and other networks. The programs allow investigation of not only the global network architecture, but also its local properties, revealing key regulators and structure-functional modules. Also, the network evolution can be traced. The package has been tested with gene, biomedical co-authorship and biomedical term word networks.

Acknowledgements: The work was supported by SD RAS Multidisciplinary Projects #21 and 39.

ON A PARALLEL ALGORITHM FOR MORPHOGENE DIFFUSION-REACTION PROCESSES SIMULATION ON A 2D CELL ENSEMBLE

Troeglazova T.S.[1]*, Baishibaev D.Ja.[1], Penenko A.V.[2], Nikolaev S.V.[3], Zubairova U.[3]
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia;*
[3]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
e-mail: tatyanameda@gmail.com
**Corresponding author*

Key words: morphogenesis, diffusion-reaction model, parallel computations, discrete-analytical scheme

Motivation and Aim: Diffusion-reaction PDE systems can be used to model morphogene transport and dynamics over a plant tissue. The aim of the work is to create parallel implementation of a diffusion-reaction solver capable for high performance simulations. High performance is needed for inverse problems solution algorithms that involve multiple direct and adjoint problem evaluations.

Methods and Algorithms: A tissue is modeled with 2D cell ensemble. A system is split with respect to different processes i.e. to the reaction and diffusion parts. To do model decomposition the additive-averaged splitting scheme is used. Reaction part is implemented with explicit discrete-analytical scheme [1] which guarantees positive morphogene concentrations.

Results: The algorithm has been realized for model of structuring the stem cell niche in shoot apical meristem of Arabidopsis Thaliana [2]. A discrete-analytical scheme performance has been tested in the application to morphogenesis.

Conclusion: Parallel structure of the algorithm allows high performance simulations of morphogenesis processes.

Availability: The program is available on request from the authors.

References:

1. V.V. Penenko et al. (2014) Variational Approach and Euler's Integrating Factors for Environmental Studies, accepted in CAMWA, 2014
2. S.V. Nikolaev et al. (2013) Model of Structuring the Stem Cell Niche in Shoot Apical Meristem of Arabidopsis thaliana, Doklady Biological Sciences, 452:336–338.

COMPUTER SIMULATION OF SELF-ORGANIZATION IN THE BACTERIAL MinCDE SYSTEM

Vitvitsky A.A.
Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia
e-mail: vitvit@ssd.sscc.ru

Key words: computer simulation. MinCDE system, cellular automata, self-organization

Motivation and Aim: Bacterial cell division begins with formation of a ring-like structure, called the Z-ring, on the cell membrane. A proper position of Z-ring in the midcell is controlled by the certain self-organization mechanisms. A bright example of these mechanisms is a MinCDE protein complex. Currently, the processes leading to a self-organization in this MinCDE system are not quite clear, but are intensively studied [1, 2]. However, exact theoretical description of these processes is a hard task because of the difficulty to obtain information about spatiotemporal dynamics of individual particles from bulk biochemical assays. Hence, computer simulation plays an important role in this problem and helps to conform (or to disprove) the proposed hypotheses.

Methods and Algorithms: Cellular Automata (CA) was chosen as a simulation tool. CA is a discrete mathematical model consisting of a set of finite state automata called cells [3]. CA models have the following advantages: ability to simulate complex nonlinear processes in active environments, simple rules of simulating and the natural fine-grained parallelism.

Results: CA-model of self-organization MinCDE protein complex based on theoretical description from [2] has been developed. The dependence of protein concentration in time and space, obtained as a result of computer simulation, revealed similarity with that, obtained by the experiments *in vitro.* In addition, the visualization of computational experiments showed propagating protein waves similar to those that emerge *in vitro*.

Conclusion: The evidence in favor of the hypothesis from [2], which claims that self-organization in the MinCDE system arises from an interplay of two opposing mechanisms: cooperative binding of MinD to the membrane, and accelerated MinD detachment due to persistent MinE rebinding have been obtained.

References:

1. M. Loose, E. Fischer-Friedrich, C. Herold, K. Kruse, P. Schwille (2011). Min protein patterns emerge from rapid rebinding and membrane interaction of MinE. Nat. Struct. Mol. Biol. 18, 577583. doi:10.1038/nsmb.2037.
2. V. Ivanov, K. Mizuuchi (2010). Multiple modes of interconverting dynamic pattern formation by bacterial cell division proteins. Proc Natl Acad Sci USA 107: 80718078.4
3. O. Bandman Implementation of Large-Scale Cellular Automata Models on Multi-Core Computers and Clusters // High Performance Computing and Simulation (HPCS), 2013 Int. Conf. IEEE Conference Publications, 2013. pp.304 – 310.

Author index

Научное издание

Тезисы VIII Международной конференции «Биоинформатика регуляции и структуры генома\ Системная биология»

на английском языке

Abstracts of the Eighth International Conference on Bioinformatics of Genome Regulation and Structure\ Systems Biology

Abstracts have been printed without editing
as received from the authors

Подготовлено к печати
в редакционно-издательском отделе
ИЦиГ СО РАН
630090, Новосибирск, пр. акад. М.А. Лаврентьева, 10

Дизайн и компьютерная верстка: А.В. Харкевич

Подписано к печати 8. 06. 2012 г.
Формат бумаги 70 × 108 1/16. Печ. л. 32. Уч.-изд. л. 36,5
Тираж 200. Заказ 245

Отпечатано в типографии ФГУП «Издательство СО РАН»
630090, Новосибирск, Морской пр., 2

MM-HPC-BBB-2014