

JEDLIK LABORATORIES REPORTS

FACULTY OF INFORMATION TECHNOLOGY AND
BIONICS AND THE JEDLIK LABORATORIES

This special issue was supported by the European Union and co-financed by the European Social Fund through the project 'Thematic Research Cooperation Establishing Innovative Informatic and Info-communication Solutions, grant number: EFOP-3.6.2-16-2017-00013.'

**Faculty of Information Technology and Bionics
and
The Jedlik Laboratories**

JEDLIK LABORATORIES REPORTS

**Vol. VII, No. 1
JLR – 1 / 2019**

EDITOR: PETER SZOLGAY

EDITORIAL BOARD:

Marco Gilli (Torino), György Karmos (Budapest), Maciej Ogorzalek (Krakow),
Sándor Pongor (Budapest), Gábor Prószéky (Budapest), Ronald Tetzlaff (Dresden)

GUEST EDITORS:

Attila Csikász- Nagy (Budapest), Gábor Szederkényi (Budapest)

Bioinspired analysis of dynamical systems and Protein network analysis

This special issue of the JEDLIK LABORATORIES REPORTS contains the works presented on the two conferences 'Bioinspired analysis of dynamical systems' and 'Protein network analysis' held jointly between 10-12 January 2019 in Esztergom, Hungary.

The conferences were supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications)

JLR – 1 /2019



PÁZMÁNY UNIVERSITY ePRESS
BUDAPEST - ESZTERGOM, 2019

© PPKE Információs Technológiai és Bionikai Kar, 2019

ISSN 2064-3942

Kiadja a Pázmány Egyetem eKiadó

Felelős kiadó

A. R. D. Szuromi Szabolcs Anzelm O. Praem.
a Pázmány Péter Katolikus Egyetem rektora

This special issue was supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

Ez a kiadvány az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg (EFOP-3.6.2-16-2017-00013, Innovatív Informatikai és Infokommunikációs Megoldásokat Megalapozó Tematikus Kutatási Együttműködések).

Bioinspired analysis of dynamical systems and Protein network analysis

joint conference
10-12 January 2019
Esztergom, Hungary

10 January 2019

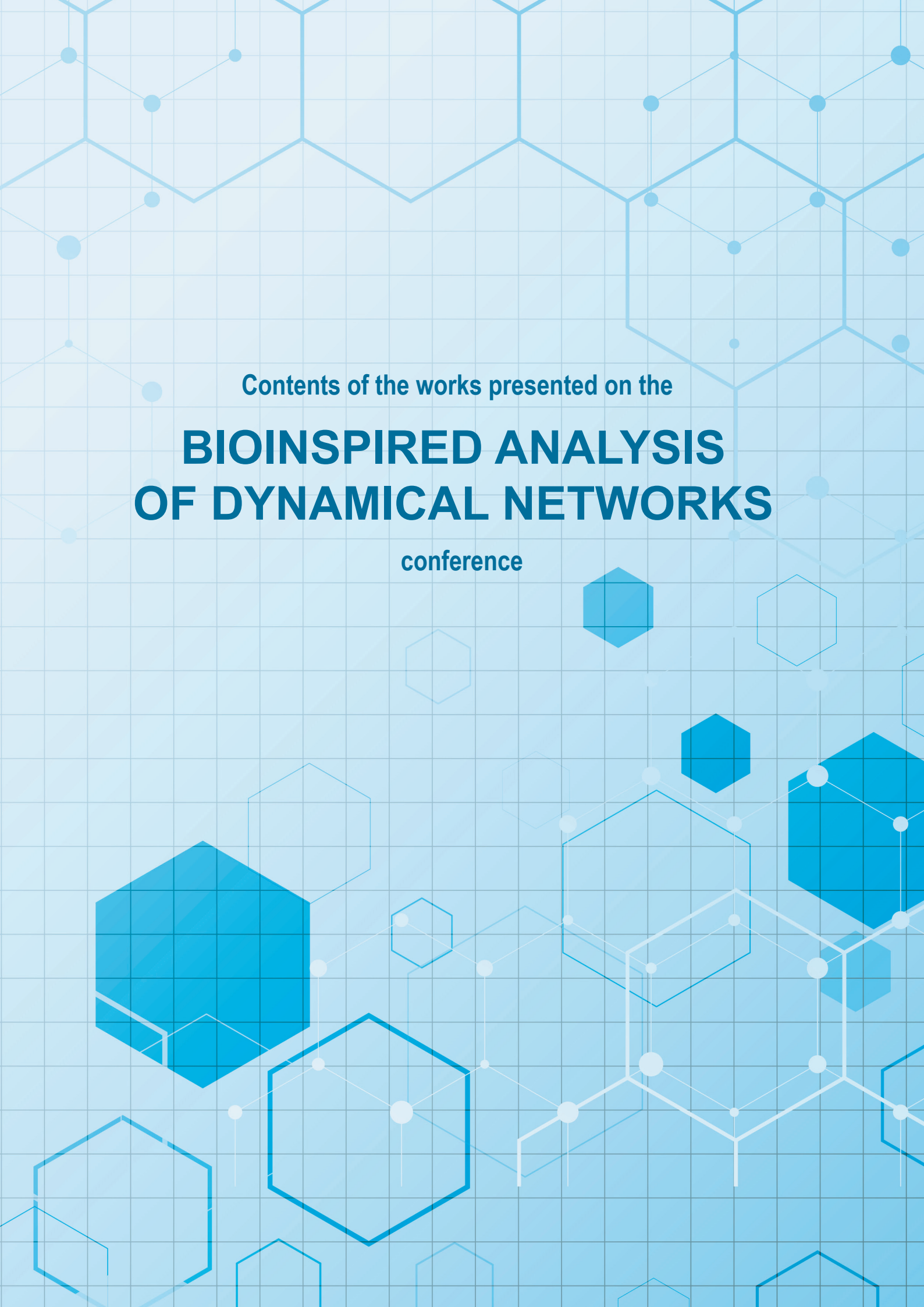
Time	Bioinspired analysis of dynamical systems	Protein network analysis
12:00 – 14:00	Registration	
14:00 – 16:00	Luca Cardelli, presented by Attila Csikász-Nagy Peter Erdi, presented by Gábor Szederkényi	
16:00 – 18:00	Poster session	
19:00 –	Gala dinner	

11 January 2019

Time	Bioinspired analysis of dynamical systems	Protein network analysis
	Chair: Kovács Levente	
9:00 – 10:00	Neil Dalchau	
10:00 – 11:00	Jörg Menche	
11:00 – 11:30	Coffee break	
	Chair: Csercsik Dávid	Chair: Gáspári Zoltán
11:30 – 13:00	Attila Répai	Csaba Kerepesi
11:30 – 13:00	Péter Battyányi	Bence Keömley-Horváth
11:30 – 13:00	Peter Szmolyan	Andrea Ciliberto
13:00 – 14:30	Lunch	
	Chair: János Tóth	Chair: Balázs Ligeti
14:30 – 16:00	Erzsébet Csuha-Jarjű	Dániel Drexler
14:30 – 16:00	Tamás Gönczy	Krisztian Buza
14:30 – 16:00	Brigitta Dűzs	Annamária Kiss-Tóth
16:00 – 16:30	Coffee break	
	Chair: Erzsébet Csuha-Jarjű	
16:30 – 17:30	Irene Otero-Muras	
17:30 – 18:30	Natasa Przulj	
19:00 –	Social dinner	

12 January 2019

Time	Bioinspired analysis of dynamical systems	Protein network analysis
	Chair: Gábor Lente	
9:00 – 10:00	Tóth János	
10:00 – 11:00	Alberto Paccanaro	
11:00 – 11:30	Coffee break	
	Chair: György Vaszil	Chair: Krisztián Buza
11:30 – 13:00	Benedek Tasi	Gábor Lente
11:30 – 13:00	Gregory Szép	Anett Hinsenkamp
11:30 – 13:00	Luca Laurenti	Gergely Szlobodnyik
13:00 – 14:30	Lunch	
14:30 – 15:30	Lars Juhl Jensen, presented by Attila Csikász-Nagy	
15:30 – 16:30	Marta Kwiatowska, presented by Gábor Szederkényi	
16:30 –	Closing and farewell	



Contents of the works presented on the

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

conference

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Keynote lecture

Dynamical systems perspective in neuroscience - historical and current approaches

Érdi, P

Center for Complex System Studies, Kalamazoo College, Kalamazoo, MI and Wigner Research Centre for Physics, Hungarian Academy of Sciences, Budapest

The brain is a prototype of hierarchical dynamical systems. The first part of the lecture reviews the basic dynamical models of single cells, networks and neural populations. In the second half it will be argued that a large class of neurological and psychiatric disorders can be understood as dynamical diseases. Specifically, a chemical kinetic model helps us to uncover some hidden links between epilepsy and Alzheimer's Disease.

PE thanks the Henry Luce Foundation for letting him serve as a Henry R. Luce professor.

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Keynote lecture

Computing Reliably with Molecular Walkers

Marta Kwiatkowska

Department of Computer Science, University of Oxford, Oxford, UK

DNA computing is emerging as a versatile technology that promises a vast range of applications, including biosensing, drug delivery and synthetic biology. DNA logic circuits can be achieved in solution using strand displacement reactions, or by decision-making molecular robots-so called 'walkers'-that traverse tracks placed on DNA 'origami' tiles.

Similarly to conventional silicon technologies, ensuring fault-free DNA circuit designs is challenging, with the difficulty compounded by the inherent unreliability of the DNA technology and lack of scientific understanding. This lecture will give an overview of computational models that capture DNA walker computation and demonstrate the role of quantitative verification and synthesis in ensuring the reliability of such systems. Future research challenges will also be discussed.

<http://www.cs.ox.ac.uk/marta.kwiatkowska/>

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Invited lecture

Model-based design of synthetic biological pattern formation

Neil Dalchau

Microsoft Research, Cambridge, UK

The Turing patterning mechanism is believed to underly the formation of repetitive structures in development, such as zebrafish stripes and mammalian digits, but it has proved difficult to isolate the specific biochemical species responsible for pattern formation. Meanwhile, synthetic biologists have designed Turing systems for implementation in cell colonies, but none have yet led to visible patterns in the laboratory.

In this talk, I will introduce our progress towards engineering synthetic gene circuits that when inserted into bacteria, enable pattern formation based on the Turing mechanism. We have explored several designs, developed new theoretical tools, and characterized a variety of biological components that can be composed into candidate networks. Our designs are also built upon the capability of using experimental measurements to quantify the properties of biological components, which enables the in silico optimisation of a target behaviour. Finally, I will introduce pattern forming circuits that rely on mutual inhibition, and mathematical analysis that enables the interpretation of stable boundary formation in such circuits. In particular, we propose a new general schema for stable boundary formation which is based on two signals with opposing gradients.

<https://www.microsoft.com/en-us/research/people/ndalchau/>

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Invited lecture

Engineering complexity in biochemical networks: methods and tools for the automated design of biochemical bistable switches and limit cycle oscillators

Irene Otero-Muras

Bioprocess Engineering Group, IIM-CSIC, Spanish National Research Council, 36208 Vigo, Spain

Biochemical network models describing the dynamics of cell signaling and gene regulation are central to systems and synthetic biology. The long term dynamics of biochemical networks, related in this context to biological function, have been subject of extensive research, in particular concerning conditions for uniqueness and stability of the equilibrium.

Here we investigate the conditions under which complex nonlinear behaviour (including limit cycles and bistability) can occur in biochemical systems from three different approaches: bifurcation theory, thermodynamics and structural analysis (based on properties of the network graph), and provide conditions for multistability and limit cycle oscillations in biochemical systems.

Our aim is not only to understand but to engineer complexity in biochemical networks. Combining the aforementioned conditions with advanced mixed integer nonlinear optimization methods (that allow scanning structure and parameter spaces simultaneously) we present here a framework for automated design of biochemical systems with predefined target behavior, with focus on bistable switches and limit cycle oscillators.

<https://sites.google.com/view/oteromuras>

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Invited lecture

Automatic model selection, automatic initial estimates

János Tóth¹, Tibor Nagy², Tamás Ladics³

¹Budapest University of Technology and Economics, Department of Analysis, jtoth@math.bme.hu

²Hungarian Academy of Sciences, Research Centre for Natural Sciences, Institute of Material and Environmental Chemistry

³John von Neumann University, Department of Science and Engineering

Given the measurements of species concentrations in a chemical reaction it is the fundamental problem is to create a mass action type kinetic model and find its parameters (most often: reaction rate coefficients or Arrhenius parameters). We propose a method for automatic model selection and a method for finding good initial estimates for the parameters, again automatically.

Automatic model construction and selection

At first, we fix the number of species (this may be the number of species measured or larger). Next, we generate all the combinatorically conceivable and chemically acceptable reaction steps (e.g. those with a kinetic order not higher than two, conserving mass, reversible, different from $X \rightarrow X$, or $2X \rightarrow 2X$, $2X \rightarrow 2Y$, etc.). If the number of species is 2, then the following reaction **steps** remain:

$\{X \leftrightarrow Y, X \leftrightarrow 2Y, Y \leftrightarrow 2X, 2X \leftrightarrow X + Y, 2Y \leftrightarrow X + Y\}$.

Then, we form **sets of reaction steps** which we call reactions. From these five steps one can form three two-step reactions if mass conservation is also required.

$\{\{X \leftrightarrow Y, 2X \leftrightarrow X + Y\}, \{X \leftrightarrow Y, 2Y \leftrightarrow X + Y\}, \{2X \leftrightarrow X + Y, 2Y \leftrightarrow X + Y\}\}$.

Now assuming constant temperature and volume and mass action type kinetics we fit all the obtained reactions (one after another) to the measured concentrations and compare the results using some measure of goodness of fit, like the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), or other, and present the best fitting reaction to the chemist for further investigations.

If detailed balance is also required, as usual, then this may imply relationships between the reaction rate coefficients, thus some of them can be calculated from these equalities. The method can also be used when not the concentrations themselves but a function or a functional of them (called **feature**, e.g. period in an oscillatory reaction) is measured.

Automatically calculated initial estimates

Models linear in the parameters (even nonlinear in the state variables) are easy to fit to measured data. The real problem starts when the model is nonlinear in the parameters as e.g. the sum of two exponentials, or the solution of a (no matter how simple) differential equation. (Cases when the model is linearizable—like the function $t \rightarrow Ae^{Bt}$ —are not really better, classical linearization methods even mislead the researcher.) The success of the art (and not pure science) of parameter estimation in the nonlinear case heavily depends on the initial estimate of the parameters. (Chemically realistic bounds are also of help, although technically they may mean a nuisance.)

To get a good initial estimate one can utilize the intuitively clear fact (what can formally be proven) that although the concentration vs. time curves in kinetic models are almost never linear functions of the parameters, but the right hand side of the induced kinetic differential equation of a reaction is [1,2,3]. The estimating process

then goes on from the initial estimate obtained using this speciality of the right hand side. We show on a few characteristic examples (Lotka-Volterra reaction without diffusion, Fisher equation with diffusion, Belousov-Zhabotinsky reaction with diffusion) with **simulated** data how this fact can be used to start from a good initial estimate and arrive at a good final estimate. The method can be extended to cases when only some of the parameters influence linearly the right hand side, or when we have compound parameters (e.g. Arrhenius form in non-isothermal systems) [5].

Methods

All the calculations from combinatorics to estimations are based on our program package ReactionKinetics [4,6] written in the Wolfram language (Mathematica) that can freely be downloaded from <http://extras.springer.com>. The codes used to the present talk can be requested from the authors.

Acknowledgements

The present work has been supported by the grant SNN 125739.

References

- [1] Halász, G.; Hangos, K. M.; Tóth, J.: Energy-optimal operation conditions of a tunnel kiln, Computers and Chemical Engineering **1988** 12 (2–3), 183–187.
- [2] Hangos, K. M.; Tóth, J.: Maximum likelihood estimation of reaction-rate constants, Computers and Chemical Engineering **1988** 12 (2–3), 135–139.
- [3] Kovács, B.; Tóth, J.: Estimating reaction rate constants with neural networks, Enformatika **2007** 26, 13–17.
- [4] Nagy, A. L.; Papp, D.; Tóth, J.: ReactionKinetics—A Mathematica package with applications, Chemical Engineering Science **2012** 83(1), 12–23.
- [5] Nagy, T.; Turányi, T.: Uncertainty of Arrhenius parameters, International Journal of Chemical Kinetics **2011** 43 (7), 359–378.
- [6] Tóth, J.; Nagy, A. L.; Papp, D.: Reaction Kinetics: Exercises, Programs and Theorems. Mathematica for Deterministic and Stochastic Kinetics, Springer Nature, **2018**, Berlin, Heidelberg, New York. (ISBN 978-1-4939-8643-9)

<http://math.bme.hu/~jtoth/index.html>

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Continuous Blood Pressure Measurement with Camera-based Optical Methods

Attila Répai^{1,3,*}, Péter Sótónyi², György Cserey¹

¹Pázmány Péter Katolikus Egyetem, Információs Technológiai és Bionikai Kar

²Városmajori Szív- és Érgyógyászati Klinika

³3in Kutatócsoport, Információs Technológiai és Bionikai Kar, Pázmány Péter Katolikus Egyetem, Esztergom

*E-mail: repai.attila@hallgato.ppke.hu

Keywords: blood pressure, pulse, non-invasive measurement

Continuously measured blood pressure (BP) signal is a crucial source of patient monitoring and diagnosis in intensive care units (ICU). Today the most widespread method is invasive measurement carried out in the lower arm arteries. As this way BP is measured directly, this provides the most accurate results, therefore it is accepted as a baseline for comparison in case of testing novel techniques. However, long term invasive measurement comes with serious risks, which motivates researchers to come up with a similarly effective non-invasive method. Here we propose a sensor design based on palpation pulse detection and we evaluate it compared to another non-invasive technique developed in our laboratory.

Introduction

In case of non-invasive BP measurement, we can sense a secondary effect of heart contractions: periodic increase in the volume of blood vessels due to the increase of arterial pressure (plethysmography), changes of the blood vessel tone (tonometry) or changes in oxygen saturation of blood (pulse oxymetry).

Inspired by the traditional manual pulse palpation technique, the sensor presented here is based on tonometry. For a such measurement to be successful, the sensor is to be positioned over an artery, at a location where there is a bone under the blood vessel and the artery is relatively close to the surface to achieve minimal attenuation caused by muscle, fat and skin tissues. Such locations are above the ulnar and radial artery at the wrist.

During the design phase the following requirements were taken into consideration.

- 1) The ability of real time, continuous, long term BP recording. In ICU, patients after complicated vascular surgeries are to be continuously monitored for several weeks. During this period, the measured BP signal should be immediately displayed to the physicians and/or stored for automated pattern recognition to help setting up a diagnose.
- 2) Easily handleable size. Current invasive measurement systems require large space due to their pipe network and fluid tanks while an optimal solution should allow more access to the patient's body. A small device is also easier to use in ambulance cars where space is even more limited.
- 3) Ease of usage and low production cost to make the device widely accessible.

Implementation

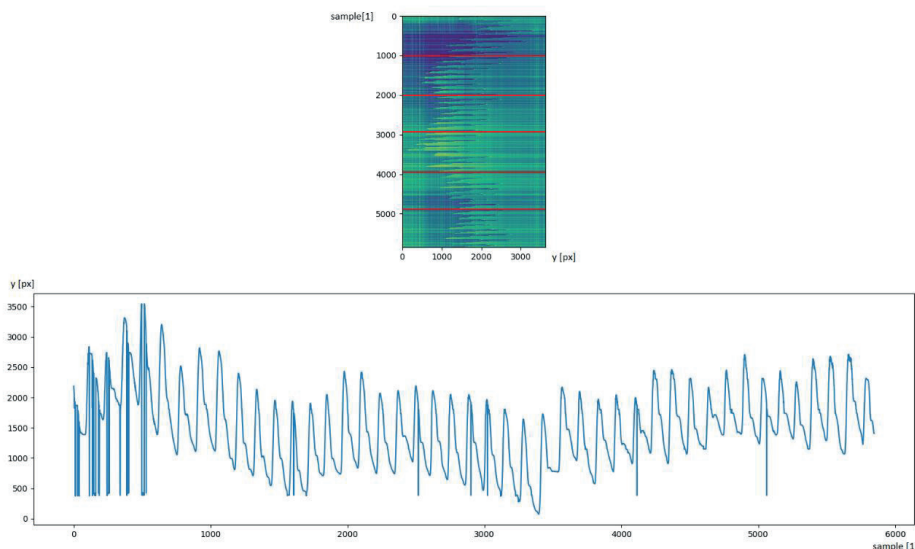
Our device consists of three main parts. The palpation unit translates the movement of the artery wall to a more observable change. The optical unit perceives this amplified change and converts it to numeric measurement data. The processing unit can store the signal, transmit it to other devices e.g. for visualization or it can perform computations such as pattern recognition.

The palpation unit contains a flexible membrane that contacts with the patient's wrist and transmits the pulsation of the artery wall to a liquid stored in a small tank. The material and shape of this membrane must fulfil three requirements: first, it has to be flexible enough not to attenuate the pulsation significantly, second, to be rigid enough to apply pressure to the blood vessel to make pulsation observable, finally, since it is in direct contact with skin, not to trigger allergic reactions.

Once the membrane is pressed by the artery with increasing volume, it displaces fluid from the tank. This fluid is lead to a pipe which has a section of low diameter and translucent material. In this section, displaced fluid causes a well detectable shift of level, which is perceived by the optical unit. This fluid has to be non-compressible and completely free from bubbles. In the experiments discussed here we used distilled water, but it showed to be suboptimal. After a period of some days, a small quantity of water leaves the pipe via diffusion, resulting in bubbles. Biological pollution of the water or the pipes can lead to the presence of algae that affect inner pipe diameter and change hydrophobicity of the pipe.

As for the optical measurement unit, we implemented two different solutions. The first uses an 5 megapixel RGB sensor (OmniVision OV5647) without infrared filter, built in the Pi NoIR evaluation board. This camera is connected to a Raspberry Pi with a high-speed CSI bus. In order to exploit the full sensor area, the built-in lens was extended with an external lens system. The internal chip of the Pi NoIR camera automatically adjusts exposition time which has to be disabled programmatically to achieve a uniform sampling rate. The maximum sampling rate is $f_s = 80$ [Hz] and the maximum amplitude is $y_{\max} - y_{\min} \approx 50$ [px]. An image processing algorithm then extracts the fluid level value from the image.

The second implementation involves a Toshiba 3648 px linear CCD sensor. This enables faster image processing which is run on an ST Nucleo microcontroller. The linear CCD offers two magnitudes higher amplitude resolution, $y_{\max} - y_{\min} \approx 1500$ [px] and a sampling frequency $f_s = 150$ [Hz]. The detected fluid level is transmitted to the host computer via SPI protocol. For the evaluation of the fluid level detection the sensor module also can be configured to send the full image to the host (in this case sampling frequency is lower due to the increased data



transmission cost). An example recording is shown in Fig. 1.

Figure 1. Concatenated image series from the linear CCD (top). Fluid level gives signal shape (bottom).

Measurement process and evaluation

The steps of measurement are the following:

- 1) The patient places the left hand on a table, hand movement is avoided.
- 2) The palpation unit is placed over the radial artery about 3 cm from the wrist, the axis of the sensor surface and the artery is parallel.
- 3) The sensor is pressed to the arm until the fluid oscillation is well observable.
- 4) The fluid level is calibrated using a screw so that the range of oscillation be in the range of the image sensor.
- 5) The recording software is started. We plan automatic base level detection to achieve automatic calibration whenever the fluid level leaves the imaging area.

For this evaluation we recorded 2 min time series.

We used recordings produced with an OptoForce OMD-20-SE-40N 3D force sensor as a reference, based on the research of Földi et al. [1].

First, we evaluated image processing algorithms, that is the formation of the pressure time series from the recorded images. In the cases of both 2D and linear CCD, the quality of level detection highly depends on the intensity of light. For the optimal quality we must ensure constant light intensity both temporally and spatially. The uncertainty of level detection can be further reduced by using hydrophobic pipe materials and non-transparent liquid. Knowing the biological model of the BP signal, single detection errors can be easily eliminated with a lowpass filter. Because of the optical measurement, this system is not sensitive to the noise originating from the power line frequency (50 Hz). Linear CCD seemed to be more sensitive to vibration, since it could cause the movement of the fluid pipe which made the oscillation out of sensing. This could be eliminated with a less elastic pipe and a firm packaging.

Second, we evaluated signal quality. We compared the signal shape properties of the measured signal to those of the theoretical model. Features used in diagnostic pattern recognition (percussion wave, tidal wave and dirotic wave, discussed in [2]) can be identified on our signal too.

Further work

To make the measured BP signal comparable to the existing BP signal databases, we need to find the correspondence between the measured oscillation amplitudes and the real pressure values. We also plan to implement automatic calibration to facilitate usage.

Acknowledgement:

The research has been supported by the European Union, co-financed by the European Social Fund(EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] S. Földi, T. Horváth, F. Zieger, P. Sótonyi and Gy. Cserey. "A novel non-invasive blood pressure waveform measuring system compared to millar applanation tonometry", Journal of Clinical Monitoring and Computing, vol. 32, 10 2017.
- [2] S. Földi. "Application areas of continuous non-invasive blood pressure measurements," PPKE PhD Proceedings, pp. 119–122, 2016.

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Programming by Multiset Transformations and the Membrane Computational Paradigm

Péter Battyányi*, György Vaszil
Faculty of Informatics, University of Debrecen
Kassai út 26, 4028 Debrecen, Hungary
E-mail*: battyanyi.peter@inf.unideb.hu

Keywords: Abstract computational models, chemical computing paradigm, membrane systems

1 Introduction

It has been taken for granted for a long time that writing parallel programs is a difficult task. Since parallelism is understood in the framework of the imperative programming paradigm, it has been proposed by Banâtre and Le Métayer to introduce a new programming style, the Gamma programming, which eludes the artificial sequentiality lying behind the imperative programming ([4]). A Gamma program is an instruction for multiset transformations, and is based on the chemical metaphor.

The idea was developed further into the notion of the chemical abstract machine ([7]) and then a formal rewriting system, the γ -calculus emerged ([3]). The concepts of the chemical abstract machine the way for the membrane calculus invented by Păun, where a structure of embedded membranes appears ([9]).

Membrane systems consist of hierarchically embedded membranes, which enclose multisets of symbolic objects. The evolution rules assigned to the regions are responsible for the development of the system. For details on the concept of membrane computing one can consult the monograph [10] or the handbook [11].

In this talk, we give the definitions of the underlying concepts concerning the membrane systems, the Gamma programming and the γ -calculus and then we establish a connection between membrane computing and the chemical computational paradigm.

Then, based on the results of [5] and [6] we present some ideas on how the chemical formalisms and membrane systems can be related to each other.

2 The Chemical Computational Paradigm

Several attempts have been made to establish a programming language suitable for programming massively parallel systems. The Gamma-formalism was invented with the aim of providing the programmer with a high-level programming language deprived of all artificial sequentiality, in which the parallel aspect is left implicit.

The multiset itself can be seen as a chemical solution, the reaction condition, R , is the property that must be satisfied so that the reaction can take place and the action A describes the effect of the reaction. The computation terminates when a stable state is reached, that is, the reaction condition cannot be satisfied

anymore. Instead of giving the exact definitions we demonstrate by an example the effect of the Γ -operation. Example 1 The Gamma-program below (reproduced from [4]) computes the maximal element of a multiset of integers M:

$$\begin{aligned} \text{maxset}(M) &= \Gamma(R;A)(M) \text{ where} \\ R(x; y) &= (x \leq y); \\ A(x; y) &= (y): \end{aligned}$$

3 The γ -calculus

The Γ -formalism has been the source of inspiration for the higher-order chemical model of Banatre et al., the so-called γ -calculus. The following presentation is based on [3].

We define molecules and patterns simultaneously in the following definition.

Definition 1 *The molecules, denoted by M , reaction conditions, denoted by C , and patterns, denoted by P , are defined as*

$$M ::= x \mid (M_1, M_2) \mid \langle M \rangle \mid \gamma(P)[C].M$$

where x is a variable standing for any molecule, (M_1, M_2) is a compound molecule built with the commutative and associative “,” constructor operator; $\gamma(P)[C].M$ is a γ -abstraction with pattern P , reaction condition C , result M , and $\langle M \rangle$ is called a solution;

$$C ::= R \mid \neg C \mid C_1 \wedge C_2 \mid C_1 \vee C_2,$$

with $R \in \{R_1, \dots, R_k\}$ where R_i , $1 \leq i \leq k$, are from a given finite set of predicate symbols of the form $R_i(\gamma_1, \dots, \gamma_{k_i})$ with $\gamma_j \in M$, $1 \leq j \leq k_i$; and

$$P ::= x \mid (P_1, P_2) \mid \langle P \rangle,$$

where x matches any molecule, (P_1, P_2) matches a compound molecule, and $\langle P \rangle$ matches an inert solution.

A solution is termed inert, if reactions can only take place inside its contained solutions. The result of a match is an assignment of molecules to variables. The first argument of match is a pattern, the second one is a molecule, its value is a substitution. The operator compares the structures of its two arguments and returns a substitution that, when applied to the pattern argument, returns the molecule of the second argument. We omit the formal details. Now we define the main rule of the calculus, the reaction rule.

Definition 2 *Let*

$$\gamma(P)[C].M, N \rightarrow \phi M, \quad (\gamma)$$

where $\text{match}(P, N) = \phi$ and ϕ assigns values to variables in such a way that $\phi(C)$ is true.

Example 2 *The γ -term $\gamma(x, y)[x \leq y].(y), (3, 4)$ reduces to (4) since in order for $\phi(x \leq y)$ to hold, $\text{match}((x, y), (3, 4)) = \phi$ should be such, that $\phi = \{x \mapsto 3, y \mapsto 4\}$. This means that $\phi(y) = 4$, thus, the result of the reaction $\gamma(x, y)[x \leq y].(y), (3, 4)$ is the molecule (4) .*

Besides the γ -rule some additional rules are also applied, for example the associativity and commutativity of the multiset constructor, which is the formalization of the Brownian motion of molecules.

4 Membrane Systems

Similarly to chemical programs, membrane systems (also called P systems) compute with multisets of symbolic objects. We mention two types of P systems here: the symbol-object P systems and the P systems with active membranes. The symbol-object P system is a structured set of regions, each containing a multiset of objects and a set of evolution rules which define the objects produced inside the system. The computation is synchronized and maximally parallel inside each membrane. The result is a number, the number of objects in a membrane labeled as output, or a multiset constituted by the number of the different objects in the output membrane.

A P system of degree $n \geq 1$ is a construct

$$\Pi = (O, \mu, w_1, \dots, w_n, R_1, \dots, R_n, out)$$

- O is an alphabet of objects,
- μ is a membrane structure of the n membranes. The out-most membrane is called the skin membrane.
- $w_i \in \mathcal{M}(O)$, $1 \leq i \leq n$, are the initial contents of the n regions.
- R_i , $1 \leq i \leq n$, are the sets of evolution or communication rules associated to the regions, and
- $out \in \{1, \dots, n\}$ is the label of the output membrane.

Example 3 Let $\Pi = (O, \mu, a^2, \emptyset, R_1, R_2, 2)$ be a system with two regions, where $\mu = [\]_2^1$ and $R_1 = \{a \rightarrow a(b, in_2)(c, in_2)^2, a^2 \rightarrow (a, out)^2\}$, $R_2 = \emptyset$. The result is the number of objects collected in region 2, which is $\{6n \mid n \geq 0\}$. The result can also be looked at as the set of vectors

$$\{(2n, 4n) \mid n \geq 0\}.$$

As mentioned above, we consider two types of P systems in this paper. In the case of *active P systems*, the rules can form the structure of the membrane system itself, moreover, we use two kinds of parities ($\{+, -\}$). We define a P system for each chemical term individually. Instead of giving the details of our construction, we present below a general definition for active P systems taken from [2].

Definition 3 Let Π be a membrane structure defined as above; the set R of development rules contains rules of the following types (for $i, j \in H, u, v, w \in V^*, \alpha, \beta, \gamma \in \{0, +, -\}$):

1. $[_i u]_i^\alpha - v[_i]_i^\beta$ (filter - out)
2. $u[_i]_i^\alpha - [_i v]_i^\beta$ (filter - in)
3. $u[_i]_i^\alpha - v$ (membrane dissolving)
4. $u[_i]_i^\alpha - [_i v]_i^\beta [_i w]_i^\gamma$ (membrane division)
5. $[_i]_i^\alpha [_j]_j^+ - [_i]_i^\beta [_j]_j^\gamma$ (subordination)

Leaning on this system of active membranes, we define a simulation of the chemical calculus by membrane systems with active membranes. This will be discussed in the next section.

5 Chemical programs and Membrane Systems

As we have seen, membrane systems and chemical programs both can be considered as representatives of models working with symbolic chemical solutions, which are multisets. In this section we state some theorems on the relation between the two paradigms.

5.1 Describing membranes systems by chemical programs

We give a short account of the simulation of membrane systems by the γ -calculus. We base our treatment on [5]. We restrict our attention to the simplest form of symbol-object membrane systems which retain the structure of the membrane system during the computations. In [1], we presented a modified translation that can also be applied to membrane systems with dissolution rules, where a membrane can dissolve in a computational step passing its objects to the parent membrane. The rules of the dissolved membrane become void.

Instead of giving the exact definitions, we mention the main idea of the translation in [5]. The objects of the membrane system are corresponded to object elements of a molecule, while the rules are translated into γ -reduction rules. Since a membrane computational step has an evolving phase, where the new elements are produced, and a communication phase where the elements newly appearing find they correct places in the membrane system. When, in a computational step, the evolving rules come to a halt, a γ -term sees to it that the elements find their correct places in the membrane, and afterwards a new computational step can begin.

5.2 Describing chemical programs by membranes systems

In order to establish the reverse direction, we define a form of membrane systems with active membranes in [6], which were obtained as a simple modification of the membrane system detailed in Section 4. For every molecule we define its unique translation as a membrane system. Then we present the rules simulating the γ -calculus reductions. We omit the details, the interested reader is referred to [6].

References

- [1] B. Aman, P. Batty'anyi, G. Ciobanu, Gy. Vaszil, Simulating P systems with membrane dissolution in a chemical calculus. *Natural Computing* 15 (4) (2016), 521-532
- [2] A. Atanasiu, C. Martin-Vide, Recursive calculus with membranes. *Fundamenta Informaticae*, 49 (1-3) (2002), 45-59.
- [3] J.P. Ban'atre, P. Fradet, Y. Radenac, Generalized multisets for chemical programming. *Mathematical Structures in Computer Science* 16 (4) (2006), 557 – 580.
- [4] J.P. Ban'atre, D. LeM'etayer, Programming by multiset transformation. *Communications of the ACM* 36 (1993), 98–111.
- [5] P. Batty'anyi, Gy. Vaszil, Describing membrane computations with a chemical calculus. *Fundamenta Informaticae* 134 (2014), 39–50.
- [6] P. Batty'anyi, Gy. Vaszil, Chemical Term Reduction with Active P Systems. In: [8], 136–150.
- [7] G. Berry, G. Boudol, The chemical abstract machine. *Theoretical Computer Science* 96 (1992), 217–248.
- [8] A. Leporati, G. Rozenberg, A. Salomaa, C. Zandron, editors, *Membrane Computing. CMC 2016.. Volume 10105 of Lecture Notes in Computer Science*, Springer, Berlin Heidelberg (2016)
- [9] Gh. P'au, Computing with membranes. *Journal of Computer and System Sciences* 61 (2000), 108–143.
- [10] Gh. P'au, *Membrane Computing - An Introduction*. Springer-Verlag, Berlin 2002.
- [11] Gh. P'au, G. Rozenberg, A. Salomaa (eds): *The Oxford Handbook of Membrane Computing*, Oxford University Press (2010)

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Multiple time-scale dynamical systems in cell biology

Peter Szmolyan^{1*}, Ilona Kosiuk¹

¹Institute for Analysis and Scientific Computing, Technische Universität Wien

*szmolyan@tuwien.ac.at

Keywords: slow-fast dynamical systems, slow manifold, geometric singular perturbation theory, scaling, mathematical models, cell biology, oscillations

During the last decades, biologists have identified a wealth of molecular components and regulatory mechanisms underlying the control of cellular processes. The molecular basis and the biochemical events involved have been studied in many systems in great detail. The complexity of the obtained biological models necessitates the development of mathematical models which can be analysed by combining analytical and computational mathematical methods. We focus on deterministic dynamical models of cellular processes, which have the form of (often large) systems of ordinary differential equations based on reaction kinetics.

Dynamical systems theory provides a powerful framework for understanding such systems. Pertinent mathematical questions of biological interest are: existence and stability of equilibria, periodic oscillations, switching phenomena, and bifurcations, i.e. changes of these dynamical properties as key parameters vary. The analysis of the global dynamics of such systems is a challenging task even for systems of small or moderate size. Detailed realistic models in molecular biology, e.g. models of specific signalling pathways or models of the cell cycle, are typically too large for direct mathematical analysis. However, there is evidence from simulations that often only a small or moderate number of components or parts of these large systems play essential roles, while large parts of the system have a more passive or even negligible role.

The mathematical reason for this behaviour is the occurrence of dependent variables and parameters with vastly different orders of magnitude. Hence, some variables may vary little and can thus be treated as constants. Some parameters may have almost no effect and can be neglected. Some variables may rapidly approach a (quasi)equilibrium and can thus be slaved to other variables. Different mechanisms may dominate the dynamics in certain regions of phase- and/or parameter-space. Identifying and analysing these regimes and the resulting decompositions into subsystems is crucial in the analysis of the dynamics. In suitably scaled variables the equations in the individual regimes often have the form of a slow-fast dynamical system and the dynamics is (locally) organized by a slow invariant manifold of lower dimension.

The commonly used quasi steady state assumption (QSSA) is a prime example of this approach, leading to dimension reduction by slaving fast variables to slow variables. However, the mathematical understanding of such processes has been considerably advanced beyond QSSA and its variants, like total QSSA. By now powerful mathematical methods collectively referred to as “Geometric Singular Perturbation Theory” (GSPT) have been developed that go considerably beyond this traditional approach [1,2,6]. These methods have been successfully used to analyze complicated dynamical phenomena in many areas of mathematical physiology, e.g. mathematical neuroscience and calcium signaling. However, slow-fast analysis and GSPT of mathematical

models of signalling pathways and the cell cycle is much less established. The main reason for this seems to be i) the more qualitative nature of available experimental data, and ii) the corresponding mathematical models typically do not have an obvious slow-fast structure. Nevertheless simulations and preliminary analysis indicate that many of these models have hidden slow-fast structures which can be utilized in the analysis. In these models a number of different scalings has to be used which are valid in different parts of phase space or parameter space. Typically, a global reduction to a single slow manifold (like in QSSA) is not possible, instead there exist several slow manifolds losing their stability at certain (more degenerate) points allowing for transitions between the various slow regimes. The simplest transitions between different slow regimes are fast jumps, well known from relaxation oscillations, but other transitions also do occur. The analysis of these transitions needed to match the different scaling regimes requires the more recently developed blow-up method, see [2,6] for an introduction and more references.

An almost universal motif found in cellular signalling networks are phosphorylation-dephosphorylation cycles generating switch-like responses and ultra-sensitivity. Cascades of such cycles form the backbone of many signalling pathways and regulatory networks. In recent work GSPT was used in the mathematical analysis of Goldbeter's model for the mitotic oscillator, which is based on a negative feedback loop containing two such cycles [3]. The key insight there was that the system has a hidden slow - fast structure caused by small Michaelis constants with several, partially degenerate slow manifolds.

Thus, arguably, slow-fast decompositions based on GSPT provide a powerful framework for successive approximations of biochemical systems starting with the fastest or dominant scales/reactions generating a sequence of approximations on slower and slower scales. The resulting decompositions of processes or models in slow and fast components, often also in additional super-slow or super-fast sub-processes, is a powerful tool for the qualitative understanding of the dynamics, responses, and bifurcations. The applicability of this approach in the analysis of the dynamics of systems based on Mass-Action, Michaelis-Menten and Hill-type kinetics has been established [3,4,7]. In ongoing work this approach is applied in the analysis of models of great biomedical importance, i.e., cell division cycle, NF-kB signalling pathway, and the MAPK- signalling pathway, see e.g. [5].

In the talk this geometric approach to multiple time scale dynamical systems is outlined and illustrated in the context of selected models from cell biology

Acknowledgments

The research of P. Szmolyan is partly funded by the Vienna Science and Technology Fund (WWTF) as project MAPKDYN under Grant MA14-049. The research of I. Kosiuk is funded as the Horizon2020 project SFSysCellBio project ID 661650.

References

- [1] N. Fenichel, Geometric singular perturbation theory, J. Differential Equations, 31 (1979), 53-98.
- [2] I. Kosiuk, P. Szmolyan. Scaling in singular perturbation problems: blowing up a relaxation oscillator. SIAM J Appl Dyn Syst, 2011, 10, 1307–1343
- [3] I. Kosiuk, P. Szmolyan. Geometric analysis of the Goldbeter minimal model for the embryonic cell cycle. J Math Biol, 2016, 72, 1337–1368
- [4] I. Kosiuk, P. Szmolyan. Switching phenomena in a model of bipolar-disorder, in preparation
- [5] I. Kosiuk, P. Szmolyan. Nonstandard relaxation oscillations in NF-kB signalling, in preparation
- [6] C. Kuehn. Multiple time scale dynamics, Springer, 2015
- [7] C. Kuehn, P. Szmolyan. Multiscale geometry of the Olsen Model and non-classical relaxation oscillations, Journal of Nonlinear Science, 2015, 25(3), 583-629

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

On Dynamics of Networks of Watson-Crick D0L Systems

Erzsébet Csuhaj-Varjú^{1*}, György Vaszil²

¹Faculty of Informatics, ELTE Eötvös Loránd University, Budapest, Hungary

²Faculty of Informatics, Debrecen University, Debrecen, Hungary

*E-mail: csuhaj@inf.elte.hu

Keywords: DNA computing, Watson-Crick complementarity, network of Watson-Crick D0L systems, recursively enumerable language, stability problems

Watson-Crick complementarity is a fundamental concept in DNA computing [11, 12], based on the corresponding biological phenomenon. The idea of a computational model where the paradigm of complementarity is considered in the operational sense was first presented in [10], by introducing the notion of a Watson-Crick D0L system.

An 0L system is a triplet $H=(V, g, w_0)$, where V is an alphabet, g is a finite substitution defined on V^* , and w_0 is an element of V^* , the axiom. A word sequence $S(H)$ of H is a sequence $w_0, w_1, w_2, \dots$ where $w_{i+1}=g(w_i)$, $i \geq 0$. The language $L(H)$ of H is the set of all words that can occur in a word sequence $S(H)$ of H . An 0L system is deterministic, a D0L system in short, if g is an endomorphism. Note that the word sequence $S(H)$ of a D0L system H is unique.

An alphabet $V=\{a_1, \dots, a_n, \bar{a}_1, \dots, \bar{a}_n\}$ is called a DNA-like alphabet. Symbols a_i and $\bar{a}_i$, $1 \leq i \leq n$, are said to be complementary symbols, we also call the non-barred symbols purines and the barred symbols pyrimidines. These terms originate from the basic DNA alphabet $\{A, G, C, T\}$ where symbols A and G are for purines and symbols C and T denote pyrimidines.

A Watson-Crick D0L system (a WD0L system) is a pair $G=(H, \Phi)$ where H is a D0L system over a DNA-like alphabet V and Φ is a mapping that defines the trigger for complementarity transition. In a DNA-like alphabet each letter has a complementary letter and this relation is symmetric. Φ is a mapping from the set of strings (words) over the DNA-like alphabet V to $\{0, 1\}$ such that the Φ -value of the axiom is 0 and whenever the Φ -value of a string is 1, then the Φ -value of its complementary string is 0. (The complementary string of a string is obtained by changing each letter to its complementary letter.)

The word sequence of a Watson-Crick D0L systems is defined as follows: when the new string has been computed by applying the homomorphism of the D0L system, then it is checked according to Φ . If the Φ -value of the obtained string is 0 (the string is a good word), then it will be the next member of the word sequence. If the obtained string is a bad one, i.e. its Φ -value is equal to 1, then the string is changed to its complementary string and that string will be the next member of the sequence. Thus, turning a string to the complement is introduced as a language theoretic operation, or in other words, the paradigm of Watson-Crick complementarity is considered in the operational sense.

A Watson-Crick D0L system G over a DNA alphabet $V=\{a_1,\dots,a_n,\bar{a}_1,\dots,\bar{a}_n\}$ is called standard if Φ is defined in such way that a word w over V is a bad word if it contains more occurrences of barred letters than non-barred letters. G is said to be with regular trigger if Φ is the characteristic function of some regular language R , that is, $\Phi(w)=0$ for a word w if it is not an element of R , otherwise $\Phi(w)=1$. Furthermore, if $\Phi(w)=1$, then $\Phi(\bar{w})=0$, where $\bar{w}$ is the complementary word of w .

Watson-Crick D0L systems and their extensions have been studied in detail: see, for example, [1,3,8,10,13,14,15,16]. Among others, it was shown that extended Watson Crick L systems are able to generate any recursively enumerable language, that is, their computational power is equal to the power of Turing machines. (An L system is extended, if the alphabet is the disjoint union of a terminal and a nonterminal alphabet where a symbol a and its complementary letter $\bar{a}$ are either terminal symbols both or nonterminal symbols both. The generated language of an extended L system contains the words that consist of terminal symbols only.) It was also shown that in the case of standard WD0L systems several problems are undecidable. One such problem is the ultimately stability problem which asks whether the generation of the word sequence of a given WD0L system contains a finite number of steps in which a turn to the complementary word takes place. Another recent development is presented in [4] where Watson-Crick T0L systems as computing devices with infinite runs are examined and related to red-green Turing machines [9]. (A T0L system – or tabled 0L system - is allowed to use different substitutions – these are called tables - in different computational steps.) Recognition by WT0L schemes (WT0L systems without axioms) is defined and it is shown that they are able to recognize (besides any recursively enumerable language) also languages that are not recursively enumerable, that is, the computational power of these types of systems is beyond the power of Turing machines.

In [5] WD0L systems were studied in a distributed framework: networks of Watson-Crick D0L systems (NWD0L systems) were introduced. Such a network is a finite collection of Watson-Crick D0L systems over the same DNA-like alphabet and with the same trigger. These WD0L systems, the components of the network, act on their own strings in a synchronized manner, and after each rewriting step, they communicate some of the obtained words to the other WD0L systems. The condition for communication is defined by the trigger for complementarity transition. In [4], two variants of communication protocols were introduced. In the case of protocol (a), after performing a rewriting step, the components keep every obtained good word and the complementary word of each obtained bad word (each corrected word), but they also send a copy of each corrected word to every other node. In the case of protocol (b), as in the previous case, the nodes keep all the good words and all the corrected ones, but communicate copies of each good string to the other nodes.

NWD0L systems have been examined from several points of view. In [4] the behaviour of particular variants of these networks using different protocols for communication was studied. In [2] it was shown that extended standard NWD0L systems form a class of computationally complete devices. An NWD0L system is extended if it is composed of extended D0L systems (ED0L systems), and it is standard if the mapping Φ defines a standard trigger for turning to the complement. The statement remains valid if incomplete communication takes place, i.e., a prefix or a suffix of the string is communicated [3]. In [6] it is shown how certain well-known NP-complete problems can be solved by using NWD0L systems in linear time. The results are obtained due to a characteristic feature of DNA computing, the massive parallelism.

In this talk we discuss problems concerning the stability of NWD0L systems. Our investigations are based on the notion of a stable word at a node and describe different languages of stable words in these systems. These concepts were introduced in [7]. A word of an NWD0L system is called stable at a node if it never leaves the node and is able to generate only itself. The generic definition of an NWD0L system is modified by distinguishing output nodes and different versions of the basic protocols, (a), (b), such that the nodes do not keep copies of the words that are communicated. The concepts of strong and weak stability languages of NWD0L systems were also introduced in [7]: The strong stability language consists of words which are stable in every output node, the weak stability language is the set of words which are stable in at least one

output node but not in all of them. We show that each word of a recursively enumerable language is an element of the strong stability language of an NWD0L system with two output nodes, and that every word of the complement language is in a weak stability language of the same NWD0L system. We also show that if certain conditions on the generation process of the words are satisfied, then equality holds instead of containment. (Since the complement of a recursively enumerable language is not necessarily recursively enumerable, the characterization of complement languages would also mean the increase of the computational power beyond the power of Turing machines.) As new development, we examine the case when the NWD0L system is given by a regular trigger. It is known that the stability problem for words of a WD0L system with regular trigger is decidable [7]. Using this technique, we describe stability languages of NWD0L systems.

Finally, we provide some open problems for future investigations.

Acknowledgment

Research supported by Grant No. K 120558 of the National Research, Development and Innovation Office, Hungary.

References

- [1] J. Csima, E. Csuhaj Varjú, A. Salomaa. Power and size of extended Watson-Crick L systems. *Theoretical Computer Science*, 2003, 290(3), 1665-1678
- [2] E. Csuhaj-Varjú. Computing by Networks of Watson-Crick D0L systems. M. Ito (Ed.) In: *Proc. Algebraic Systems, Formal Languages and Computation*, RIMS Kokyuroku1166, Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 2000, 43-51.
- [3] E. Csuhaj-Varjú. Networks of standard Watson-Crick D0L systems with incomplete information communication. J. Karhumäki et al. (Eds.) In: *Theory is Forever. Lecture Notes in Computer Science 3113*, Springer Verlag, Berlin-Heidelberg, 2004, 35-48.
- [4] E. Csuhaj-Varjú, R. Freund, Gy. Vaszil. Watson-Crick T0L Systems and Red-Green Register Machines. *Fundamenta Informaticae*, 2017, 155(1-2), 111-129
- [5] E. Csuhaj-Varjú, A. Salomaa. Networks of Watson-Crick D0L systems. M. Ito, T. Imaoka (Eds.) In: *Words, Languages & Combinatorics III. Proceedings of the International Colloquium*, Kyoto, Japan, March 14-21, 2000. World Scientific Publishing Co., Singapore, 2003, 134-149.
- [6] E. Csuhaj-Varjú, A. Salomaa. The Power of Networks of Watson-Crick D0L Systems. N. Jonoska et al. (Eds.) *Aspects of Molecular Computing. Lecture Notes in Computer Science 2950*, Springer, 2004, 106-118.
- [7] E. Csuhaj-Varjú, Gy. Vaszil. Stability Languages of Networks of Watson-Crick D0L Systems. Presentation at

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Spatio-temporal Modelling of Tumor Progression and Angiogenesis

Tamás Gönczy^{1, 2, 3, *}, Dávid Csercsik^{1, 2}, Levente Kovács²

¹ Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary

² Physiological Controls Research Center, Research and Innovation Center of Óbuda University, Budapest, Hungary

³in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: tamas.gonczy@gmail.com

Keywords: angiogenesis, cellular automaton, tumor progression

Introduction

Understanding individual and collective cell behaviour and cell-level interactions plays a crucial role in speeding up advances in cancer research and gaining deeper insight into the tissue-level pathophysiology of cancerous diseases. Basic cellular functions are greatly influenced by cell mutations and the over-balancing of different growth factors, which results in an extreme diversity in terms of tumor-specific features and progression. In order to develop effective therapies and fully understand the temporal evolution of the tumor cell population, we need to consider the type, molecular background and cell-level mechanisms of different vascularization processes.

The progression of solid tumors is highly dependent on neovascularization. Following a rapid cell proliferation and growth, in the absence of newly developed capillaries, most solid tumors enter a dormant state when reaching 2-3 mm in size [1]. In the case of solid tumor growth, angiogenesis is considered to be responsible for capillary formation, tissue invasion and the formation of metastases. The process of angiogenesis is mainly associated with hypoxia-driven sprouting of new microvascular capillaries from pre-existing postcapillary venules involving VEGF and other growth factors [2, 3]. On the other hand, we have to note that a number of researches showed that next to sprouting angiogenesis, other alternative mechanisms are also involved in the development of tumor vasculature [4].

Several studies have investigated antiangiogenic therapies and drugs, and showed that in many cases the overall effectiveness is far behind expectations. These studies stated that the main cause of the decline in efficiency is the poorly chosen drug administration strategy [5]. The need for patient-specific therapies is widely recognized in modern medication. An individualized targeted molecular therapy can greatly improve the quality of life for cancer patients, lower costs and increase the effectiveness of the treatments.

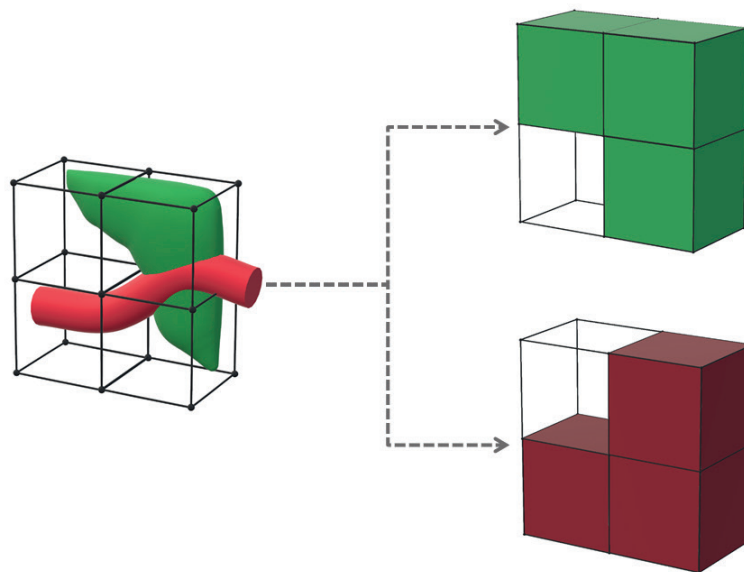
Most of the modern computational models can reliably reproduce a set of tumor related phenomena by simulating certain steps and interactions on different scales regarding time and space. On the other hand, these solutions often focus on providing a very detailed mathematical description of background processes and lack the ability of simulating higher level interactions that might be of great importance. We present a 3-dimensional hybrid cellular automaton (CA) model of tumor growth and tumor-induced angiogenesis that explicitly models different vascularization mechanisms accompanying the most critical phases of tumor progression. Additionally, the

model allows us to test different drug administration strategies under specific physiological circumstances by simulating the effects of antiangiogenic therapy and structural inhomogeneities of the host tissue.

Model Description

Processes simulated by our model include tumor cell proliferation and necrosis, tumor-induced vascularization mechanisms and the diffusion of tumor angiogenic factors (TAF), oxygen and nutrients. The proposed model is also capable of simulating the effects of antiangiogenic drug administration and structural inhomogeneities of the host tissue. While many of the existing solutions mainly focus on cellular processes and interactions that affect endothelial cells (elongation, migration, etc.), our model simulates

the behavior of cell groups on multiple levels. The proposed model describes the environment of the tumor population as a 3-dimensional grid, where each cell is represented by an entry of a simple numerical matrix. This makes optimization steps straightforward by reducing the complexity of the underlying computations and provides an adequate level of modularity. The basic functions of the 3-dimensional model are derived from a



previously developed hybrid cellular automaton model of tumor-induced angiogenesis [6].

Figure 1. In the proposed model each cell in the 3-dimensional grid is described by the state value and the concentration of nutrients, TAF and antiangiogenic inhibitor. As the figure shows, a grid cell can contain both tumor cells and capillaries at the same time.

The proposed model can be divided into four main interconnected modules. Each of these modules is represented by a 3-dimensional numerical array that holds the current state of each grid location as shown in Figure 1. In the following, let us summarize the mechanisms which are taken into account during the simulation.

Tumor growth: Tissue volumes, that are adequately supplied with nutrients provide the opportunity for tumor cells to spread further into the host tissue. This process highly depends on the supply provided by adjacent blood vessels, either initially existing ones, or capillaries developed by tumor-induced angiogenesis. Cells can survive exclusively within the diffusion range ($100\mu\text{m}$ [8]) measured from the closest capillary. Tumor cell proliferation is also limited by structural inhomogeneities of the investigated tissue compartment.

TAF diffusion: Hypoxic cells of the tumor population produce proangiogenic factors that can trigger cellular processes inside endothelial cells of the neighboring capillaries, which lead to the formation of new blood vessels. The diffusion of these substances in the interstitial space determines the dynamics of vascularization processes, and it is also the target of antiangiogenic therapies.

Blood vessel evolution: When the tumor enters the state where the phenomenon of the angiogenic switch

can be observed, the process of tumor-induced angiogenesis is initiated. During this process, diffusing angiogenic growth factors stimulate nearby capillaries to grow new branches in the direction of the stimulus. This mechanism is responsible for the development of a supplying vasculature for the increased metabolic needs of the starving tumor population.

Nutrient diffusion: The proliferation and survival of cancerous cells depends heavily on the amount of oxygen and other essential substances. This means that without a mature vascular network, the proliferation of tumor cells will be limited and the tumor can not grow further.

Continuous processes of biological systems can be described mathematically with differential equations (DE) or a system of DEs. Nevertheless the spatial descriptive capabilities of such systems are considered to be inadequate if the goal is to model the dynamics of complex biological systems. Next to the effective spatial modelling capabilities, another outstanding feature of the CA setup is that it does not describe complex systems with complex equations, but model them by describing interactions of simple components following simple rules. Ordinary differential equations can be approximated by finite difference equations (FDE), what allows us to formulate the equations of discrete state variables of the CA.

State Equations of the CA:

$$S_T(t) = f_T(S_T(t-1), P_P(t), P_N(t), P_H(t), N_T(t-1)) \quad (1)$$

$$S_V(t) = f_V(S_V(t-1), P_V(t), N_V(t-1)) \quad (2)$$

$$C_N(t) = f_D(C_N(t-1), N_{C_N}(t-1)) \quad (3)$$

$$C_{TAF}(t) = f_D(C_{TAF}(t-1), N_{C_{TAF}}(t-1)) \quad (4)$$

$$C_A(t) = f_D(C_A(t-1), N_{C_A}(t-1)) \quad (5)$$

Where $S_T(t)$ is the state of a grid cell in the tumor matrix and $S_V(t)$ is the cell state for the vasculature matrix. $C_N(t)$, $C_{TAF}(t)$ and $C_A(t)$ are the concentrations of nutrients, TAF and antiangiogenic inhibitor respectively. $N_i(t-1)$ is the notation for the Von Neumann neighborhood of the given grid cell containing all the corresponding states or concentrations of neighboring cells. Each state and concentration is updated in every iteration using transition functions (f_T , f_V , f_D). f_T denotes the state transition function for tumor cell proliferation, necrosis and hypoxia, while f_V is the state transition function of the vascularization process. f_D refers to the 3-dimensional diffusion process which is described by a discretized differential equation that defines the density of the diffusing material at the given grid cell.

Concentrations are directly connected to probability values, thus a normalized level of the diffusing substances are considered in the equations listed below. Probability values of tumor progression are calculated using the following equations:

$$P_P(t) = (1 - I_{x,y,z}) \cdot (P_P + C_N(t)) \quad (6)$$

$$P_N(t) = C_N(t) \quad (7)$$

$$P_H(t) = 1 - C_N(t) \quad (8)$$

Where $I_{x,y,z}$ refers to the inhomogeneity value at grid cell that is in the location given by indices x , y , z . The equation of probability values for different rules resulting in vascular network development can be written in the form:

$$P_V(t) = (1 - I_{x,y,z}) \cdot P_{rule} + C_{TAF} - C_A \quad (9)$$

Conclusion

The simple structure of the CA and the inclusion of discretized differential equations provides beneficial features to the proposed model. By explicitly describing alternative vascularization mechanisms, our model can reproduce characteristics of tumor vasculature observed in in vivo measurements and provides an effective simulation framework for testing and designing individualized and optimized antiangiogenic therapies.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 679681) and from the EFOP-3.6.3-VEKOP-16-2017-00002 project.

The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] Folkman, J. (1971), Tumor angiogenesis: therapeutic implications. *New england journal of medicine*, 285 (21):1182–1186.
- [2] Ricciuti, B., Foglietta, J., Bianconi, V., Sahebkar, A., and Pirro, M. (2017), Enzymes involved in tumor-driven angiogenesis: A valuable target for anticancer therapy. In *Seminars in cancer biology*. Elsevier.
- [3] Simons, M. (2005), Angiogenesis: where do we stand now? *Circulation*, 111(12):1556–1566.
- [4] Döme, B., Hendrix, M. J., Paku, S., Tóvári, J., and Tímár, J. (2007), Alternative vascularization mechanisms in cancer: Pathology and therapeutic implications. *The American journal of pathology*, 170(1):1–15.
- [5] Sápi, J., Kovács, L., Drexler, D. A., Kocsis, P., Gajári, D., and Sápi, Z. (2017), Tumor volume estimation and quasi-continuous administration for most effective bevacizumab therapy. *PloS one*, 10(11):e0142190.
- [6] Gönczy, T., Csercsik, D., and Kovács, L. (2018). A hybrid cellular automaton model of tumor-induced angiogenesis. In *2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*. IEEE.
- [7] Toffoli, Tommaso; Margolus, Norman (1987), *Cellular Automata Machines: A New Environment for Modeling*, MIT Press, p. 60.
- [8] Vaupel, P. (2004). Tumor microenvironmental physiology and its implications for radiation oncology. In *Seminars in radiation oncology*, volume 14, pages 198–206. Elsevier.

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Chemical analogues of biological periodicity: the immanent potential of pH oscillators

Brigitta Dúzs^{1,2,*}, István Szalai¹

¹Institute of Chemistry, Eötvös Loránd University, Pázmány Péter sétány 1/A, H-1117, Budapest, Hungary

*E-mail: duzsbrigitta@gmail.com

Keywords: nonlinear kinetics, periodicity, reaction-diffusion, pattern formation, pH oscillators

Temporal and spatial periodicities are basic attributes of living organisms as signaling pathways often generate oscillatory responses. Temporal oscillators drive temporal cycles, e.g. cell and circadian cycles, by controlling the periodic accumulation or activity of a biomolecule. Spatial oscillators define subcellular positions, e.g. the site of cell division, by periodic variation in the localization of a protein. Beyond the understanding of the exact molecular mechanism of the biochemical oscillators it is also challenging to explore the fundamental elements of nonlinear dynamics in these systems such as negative and positive feedback, time delay, sufficient nonlinearity of the reaction kinetics and random fluctuations (noise). The interaction of these phenomena may results in complex behaviors such as bistability, excitability or oscillations in non-equilibrium conditions, and the combination of reaction kinetics and diffusive transport leads to spatial self-organization such as fronts, travelling waves and stationary (e.g. Turing type) patterns. These universal dynamics can be developed and investigated efficiently with purely chemical reaction networks, which are considered as analogues of living systems in some aspects. To achieve this goal, we need well-designed reactor geometries to adjust the initial and boundary conditions.

Here we focus on pH oscillators [1] as the pH is a common input signal for cells. The members of the Landolt-type family of pH oscillators are good candidates to generate time and space periodic phenomena in different reactor geometries. These systems are based on the hydrogen ion autocatalytic oxidation of a substrate (e.g. the oxidation of HSO_3^- by BrO_3^-) which is coupled with a hydrogen ion consuming negative feedback process (e.g. reaction with $[\text{Fe}(\text{CN})_6]^{4-}$). The variants of the initial chemicals are advantageous in the experimental study of the general features of the systems. In a well-stirred continuously supplied reactor bistability and sustained concentration oscillations develop depending on the input parameters. The bulk oscillations appear with the amplitude of 3-4 pH units which is suitable to extend the system with a coupled pH-induced process (e.g. complexation or self-assembly) [1]. Sustained reaction-diffusion dynamics can be investigated in gel reactors where a piece of hydrogel is fed by fresh chemicals from an outer compartment. By adding a hydrogen ion binding polymer, this configuration gives rise to Turing instability [2] which is known as a paradigm of morphogenesis. Recently, we have investigated the configuration where the gel is fed from two compartments with complementary sets of initial reactants. Here the Turing pattern formation is localized by the antagonistic chemical gradients which is an analogue of Wolpert type coding of positional information in developmental biology.

Acknowledgments

The authors thank the support of the National Research, Development and Innovation Fund (119360).

References

- [1] M. Orbán, K. Kurin-Csörgei, I. R. Epstein. pH-regulated chemical oscillators. *Accounts of Chemical Research*, 2015, 48, 593-601
- [2] I. Szalai, J. Horváth, P. De Kepper. Contribution to an effective design method for stationary reaction-diffusion patterns. *Chaos*, 2015, 25, 064311

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Design and Development of an Anatomically Correct, Biomechatronic Hand Model

Benedek József Tasi^{1,4,*}, Miklós Koller², and György Cserey³

Authors are with the Faculty of Information Technology and Bionics of Pázmány Péter Catholic University, 1083 Budapest, Hungary

¹Info-bionics Engineering MSc student at the Faculty

²Senior lecturer at the Faculty

³Associate professor and vice dean of innovation at the Faculty

⁴In Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: tasibenedekjozsef@gmail.com

Keywords: bio-inspired robotics, prosthetics, biomechatronics, biomimetics, robotic hand, functional anatomy

The general aim of neuroprosthetic research is to develop devices that can be reliably used in everyday situations. Most existing prosthetic devices are designed to restore the primary abilities of the human hand, like grasping simple objects and producing basic gestures. This is usually achieved by following a conventional approach of employing mechanized joints and actuation systems with significantly reduced biomechanical complexity compared to the biological ligaments and tendinous networks. Some of the most advanced examples for this are the Shadow Dexterous Hand of the Shadow Robot Company [1] and the Biomechatronic Artificial Hand developed by T. Zhang et al. [2]. However, while the degrees of freedom and overall capabilities of these models are useful for approximating the behavior of the human hand, the biomechanical features are excessively roboticized. This makes their movements lack the natural, “organic” feeling of the biological hand, which is further enhanced by the severance of the biologically present mechanical dependency between joint positions and muscle activity, and thus significantly reduces efficiency outside the boundaries of under-actuated prosthetic or industrial applications.

We believe that with prosthetic hands, the true challenge lies in restoring not just general functionality, but the delicate dexterity and fine-motoric skills the human hand. Therefore, it is possible to improve upon these concepts by realizing the elaborate anatomical structure of the hand, which is essential for human-like dexterity. But since these features are extremely complex, and there is still no complete consensus on their exact biomechanical utility, there is only few research attempting to fully implement these in robotic models. Additionally, while the selective inclusion of the components with the most prominent functionality has proven to be sufficient for producing some grasps and gestures in a dexterous and somewhat biomimetic way, the abundance of intrinsic and extrinsic muscles and the variable geometrical features of the musculoskeletal and ligamentous system are highly critical for producing precision movements, like handwriting or playing musical instruments.

Z. Xu et al. have developed a Highly Biomimetic Anthropomorphic Hand [3], which implements the extensor network and the flexor pulleys, along with crocheted ligaments in an ergonomic design that is capable of grasping and manipulating objects in a satisfyingly human-like way. This model also successfully demonstrates the dominance of the most significant muscles of the hand, by employing certain simplifications like the passivation of the interosseous, complete exclusion of the superficial long flexor and lumbrical muscles, partial exclusion of the intrinsic and extrinsic thumb muscles and the modification of their tendinous network. Despite the omission of these subdominant features, the model can efficiently and dexterously manipulate several objects. Therefore, the impact of these omissions on dexterity might not be immediately apparent upon observing grasps and other basic functionality, but there is a critical reason for the fact that these features have not disappeared in the process of evolution. For more complex and precise movements, like handwriting, the importance of the synergistic actions of all finger-actuating muscles must be emphasized [4], and if the goal is

the complete restoration of hand functionality, not one of them should be neglected.

The Human-like Robotic hand, presented by A.A.M. Faudzi et al. [5] is a significant step in the direction mentioned above. In this model, all notable intrinsic and extrinsic muscles are implemented using thin McKibben artificial muscles, which helps in the realization of not only the biomechanical properties and capabilities of the hand, but the volume and orientation of its musculoskeletal system as well. Although novel in concept, this model was designed for research purposes, not application in industrial areas or every-day use. While this approach is helpful for examining the contribution of each muscle to different finger movements, the string-based tendinous network and the use of adhesives for fixing the ligaments in place were not designed for significant robustness, while the analog, pneumatic actuation system could lead to difficulties both in the terms of compactness, and the fitting of the prosthetic on the human body.

Our goal is somewhere in-between: the purpose of this study is to thoroughly understand and model the intricate structure of the human hand, which sources its extreme dexterity and precision, and develop a biomechatronic model which includes the interpretation of every distinguishable, functionally contributing anatomical detail like articular surface geometry, the elaborate tendinous and ligamentous system alongside every intrinsic and extrinsic muscle, while providing a stable template for the possibility of significant force output. The need is also recognized for averaging the wide range of different anatomical variations of the features occurring in the human hand, and attempting to preserve the characteristics of the connective tissue structure of the functionally independent compartments, meaning that while an engineered model will always be constructed from separately fabricated elements, in biology, everything is coadunate to a certain level, which must be accounted for.

As the anatomical diversity of the human hand implies that the functional aspects [4] are more important than the kinematics or extreme biomechanical accuracy, our study settles on a more empirical approach: Instead of designing the model primarily with the desired skillset and mechanical capabilities in mind, the direction is to accurately mimic the anatomical structure, engineer an efficient mechatronic actuation system to substitute the biological muscles, then evaluate the resulting capabilities and refine our design accordingly, as our ultimate goal is the accurate replication of hand functionality in both powerful and precise movements.

In our design, the geometrical details of all 27 hand bones are preserved through the use of anatomical 3D models generated from MRI images (provided by the BodyParts3D site of Anatomography [6]) as a template. The formation of firm grips strongly depends on the fingers' ability to passively conform to the shape of the held object. As these displacements and deformations are limited by the contacting surfaces and the joint ligaments, accurately modeling the protrusions of the bone heads and bases is essential, a feature which would be difficult to achieve using fixed-axis hinge joints that are usually employed in prosthetic research. To this end, the bones of the fingers and wrist are 3D printed from rigid Stratasys VeroWhite resin-based material. As wrist movements are out of scope of the current stage of this study, but the geometry of the carpal area is strongly contributing to the formation of tendon pathways and finger orientation, the carpal bones are fused together and fixed to the metacarpals.

Silicone of ShoreA60-80 hardness is used to create an artificial representation of the joint capsule. We have no need for a closed system to contain the synovial fluid, so only the functionally relevant, thickened fibrous bundles, the collateral ligaments are included. The primary bundle, the proper collateral ligament is responsible for tethering the bones to each other on the two sides, but we also aim to emphasize the importance of the volar plate: it not only serves as a check-rein (along with the anterior capsule recess and the check-rein ligaments) to prevent hyperextension, but also enhances joint stability and movement dynamics through ensuring maximal surface contact. Furthermore, it allows for the accurate inclusion of the accessory collateral ligaments (that suspend the volar plate from the bone head), the elasticity of which plays an important role in coordinating both flexion and extension. The 1 mm thick, laser-cut silicone strips are screwed firmly onto the bones in their anatomical insertion points, to allow for unobstructed tendon sliding over them. This joint design aims to promote our model's capability for natural movements in terms of active and passive movement range and limitations of conformation, with dynamics similar to that of the human hand, while keeping fabrication time and complexity at the minimum.

In the human hand, finger actuation is realized by a total of 37 muscles. 18 of these are extrinsic, generally larger muscles, with their bellies located in the forearm, while the other 19 are intrinsic, smaller ones, with their full volume located distally to the wrist. These can be grouped into two functionally different classes: the flexors

for flexing, and the extensors for extending the fingers. Generally, each long finger is actuated by six muscles (excluding the extra extensors of the index and little fingers, which are surgically deemed as less important): two extrinsic long flexors, the flexor digitorum profundus and sublimis, one extrinsic long extensor, the extensor digitorum communis, two intrinsic interosseous muscles, and one intrinsic lumbrical, the synergistic and partially antagonistic actions of which are responsible for coordinating joint movements.

The long flexors are coursing through the flexor retinaculum of the wrist and a series of elastic pulleys on the palmar surface of each bone, before reaching their insertions to the bases of the distal and proximal phalanges. This structure ensures optimal angle of contact between bone and tendon, while pulley elasticity results in a bulging effect, slowing finger movement and increasing gripping force with finger flexion.

The pathway of the long extensor, however, is significantly more complicated: it merges with the tendons of the interossei and lumbrical muscles to form an elaborate tendinous network, the extensor system [7], which is responsible for transmitting the torques generated by muscle activity (which serve a different, well-characterized purpose for each muscle group [8]). The geometrical features of this extensor network are the most notable implemented anatomical details, as they are the primary effectors of the co-dependency and synergy of muscle activities that effect independent flexion and extension of the metacarpophalangeal and the two interphalangeal joints, ensure identical rate of flexion and extension at the interphalangeal joints, and enhance grasping precision and strength, while maintaining a different level of tendon stiffness at each area to limit passive joint conformation.

We use laser-cut 1 mm thick silicone of ShoreA60 hardness to mimic the elastic pulley system and the flexor and extensor retinaculum, riveted to the bone shafts in multiple points. The flexor tendons are realized as 1 mm diameter shoe thread strands, which are connected to screws in their anatomical insertion points. The extensor network was carefully designed to mimic the anatomically determined geometry and the biological insertion points, with the previously described important functional characteristics in mind. It is laser-cut from silicone of 1 mm thickness and ShoreA80 hardness, wrapped around the finger, attached to the bones with screws, with thread strands attached to their loose ends to realize their entire tendon pathways. The multi-layered nature of the system is also accounted for by strands interconnecting the silicone bundles, with the “protruding” areas being realized by strings connecting the silicone network to screws sunk into the bones.

A specifically focused part of this study is the opposable thumb, as it plays a critical and unique role in the function of the hand, being essential to the formation of most grasps, and providing the hand with most of its distinguishable functionality. With respect to opposition, the importance of its carpometacarpal joint [9] must be highlighted: the joint is composed of two saddle-shaped surfaces, the inner ridges of which are transposed asymmetrically. They are held together by several elastic ligaments, the laxness of which enables considerate play, allowing the metacarpal bone to take almost any position in space that is not limited by the contacting surfaces, but effecting a slightly oblique finger trajectory by the means of active axial rotation during movement. Different degrees of displacement and axial rotation (active and passive) are possible in each direction, limited by the ligaments, which is why truly human-like movements would be difficult to achieve using a simple ball joint or generalized ligaments.

For taking advantage of the highly capable joint structure, the thumb utilizes an abundance of extrinsic and intrinsic muscles [4]. These allow for the precise independent control of the joints, and eliminate the need for elaborate extensor network geometry. The extrinsic muscles exert most of the finger’s gripping power, with a single long flexor affecting all of its three joints, which is antagonized by a long extensor at each. This system essentially allows for the individual control of each joint, since the flexor can only affect the ones which are not stabilized by their extensors.

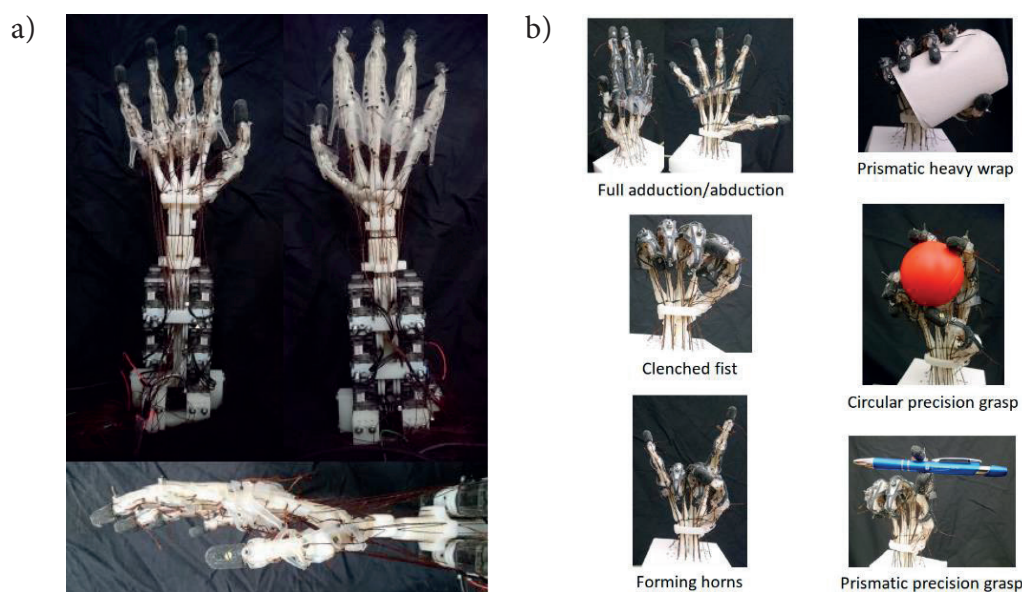
The intrinsic muscles, which can be divided into a lateral and medial group based on their location in the palm and the thenar eminence, contribute to different variations of thumb opposition and precision enhancement. This is enabled by the structure of the metacarpophalangeal joint of the thumb being slightly differs from the well-defined, 2 DoF version of the long fingers: the laxness and elasticity of its collateral ligaments allow for the displacement and axial rotation of the proximal phalanx in both directions as a result of intrinsic muscle activity, enabling the fingertip to turn towards any other digit, facilitating full opposition. The intrinsic muscles exert their torques either directly on the phalanges, or the two sesamoid bones on the anterior surface of the volar plate, which are tethered to the joint capsule by thin fibrous bundles, and smartly distribute muscle torques to different

parts of the complex. Furthermore, a few of their fibers wrap around the proximal phalanx to merge with the longest extensor in the middle and form the thumb's extensor expansion, which allows them to assist the long extensor in the extension of the interphalangeal joint.

These elaborate functionalities make it extremely important for us to accurately realize the geometry of these muscles and tendons, as the directions of their exerted forces are primarily responsible for dexterous and human-like thumb behavior by promoting the active and passive conformation of the thumb to the shape of any object, which increases grasp strength and stability. The extrinsic muscles are modelled by single thread strands, as most of them have no intertendinous connections. The pulley system of the long flexor is realized similarly to that of the long fingers, but at the level of the wrist, thumb extensors also have their own separate silicone retinacula. The extensor expansion is laser-cut from 1 mm thick silicone of ShoreA80 hardness, which is wrapped around the PP with its distal end realizing the insertion of the longest extensor. Its lateral and medial edges are fixed to the silicone joint capsule and the sesamoids, which are realized as 3D printed toruses with an outer diameter of 5 mm.

In this model, we mimic each included muscle with a separate Dynamixel smart servo motor (developed by Robotis). This allows us to examine and replicate the individual and synergistic activity of almost every hand muscle with precision control assisted by position, speed and torque feedback. The Dynamixel family provides several options, of which three different variants are used in this study: the XL-320 is a compact servo with relatively high torque (0.4 Nm) and speed (114 rpm) with a movement range of 300° (in 1024 steps), suitable for mimicking most of the muscles. The AX-12A trades some speed (59 rpm) for much larger torque (1.5 Nm) while retaining the same range of movement, facilitating its application as muscles responsible for gripping strength and stabilization. To save actuator space in the forearm-segment, a few muscles with the highest level of coactivity can even be slaved together using this servo. The MX-12W trades some torque (0.2 Nm) for precision ($4096 \text{ steps}/360^\circ$) and speed (470 rpm), and introduces the multi-turn feature, which allows the servo to employ a resolution divider to trade steps/turn for movement range. This allows us to work with smaller servo drums to increase torque output while maintaining precision and speed, making it suitable for both large movements, for which the other servos with limited range would prove insufficient, and the precise coordination of most intrinsic muscles.

All of the servos are located in a forearm-based scaffolding: the tendons (1 mm diameter thread) of the extrinsic muscles are biomimetically laced through the flexor and extensor retinacula of the wrist, while the intrinsic tendon pathways lead through tunnels in the bones at their original insertion points to preserve their anatomical geometry and exerted force directions, while simultaneously allowing for the convenience of external actuation. The vertical tendon pathways are then guided to the Dynamixels by a series of 3D printed tunnel components. It is important to point out that several of the thenar muscles are two-headed, meaning that they arise from two



different areas or bones. This greatly

Figure 1. – The complete Anatomically Correct Biomechatronic hand model, and the demonstration of its capabilities

- a) The hand model and its actuation system, palmar (top left), dorsal (top right) and radial (bottom) view
- b) Functional demonstration, basic free gestures (left) and basic prehensile hand taxonomy realization (right)

affects the direction of their exerted forces, a feature which must also be taken into consideration while designing their pathways and bone sheaths. Furthermore, as the volume and thickness of the intrinsic muscles are also strongly contributing to the formation of the tendon pathways and palm surface, we employ string guides to further imitate biological muscle span and directions. The base of the scaffolding houses an Arduino Uno microcontroller and a HC-06 Bluetooth chip, through which we can control the servos remotely, using a computer or a mobile device.

The results of the implementation of the previously detailed anatomical guidelines is shown on Figure 1/a. This anatomical hand prototype bears strong structural and ergonomical similarities to the biological hand. With the basic control interface, simple grasps and gestures of varying types can be demonstrated. A striking difference from previously examined mechanized models is the human-like, organic nature of movement execution, both regarding finger trajectories and joint rotation dynamics. The poise of the used materials and the ligamentous design allows for similarly human-like passive deformation in grips or upon applying external strain. Figure 1/b shows the exploration of the hand's basic prehensile taxonomy by performing example grasps and postures to demonstrate the overall capabilities and versatility of our model.

As of now, we have successfully designed and prototyped a competent, anatomically correct biomechatronic model of the human hand, which can be used as a template for refinement and advanced research in our future studies. During the research process, several new concepts were realized towards the improvement of the current design.

Conducting a thorough, functional cadaveric study with the assistance of leading hand surgeons will be necessary during the next step, in order to personally examine the remaining questions regarding the anatomy of the hand from a modelling point of view, and to perform measurements regarding the material qualities of the muscles, tendons and ligaments, and gain a better understanding of biological coadunation between the individual features. Studying and implementing wrist functionality is another important goal, which would be facilitated by the inclusion of artificial muscles as actuators. This would allow for more biomimetic muscle and tendon pathway realization in the forearm, and truly intrinsic implementation of the intrinsic muscles. We also intend to explore other options for usable materials and manufacturing methods to shorten fabrication time and improve robustness and user comfort.

Thorough examination of the methods of neural prosthetic control is required to develop an interface to make our model applicable as a prosthetic device. Extensive research in the field of invasive and non-invasive electrode placement, and communication protocols with the central and peripheral nervous system is required. We will explore methods that promote identification of individual motor commands (either in the form of muscle or nervous activity), preferably even in the deeper layers of the forearm and the area of the hand, in order to acquire signals that are suitable for dexterous prosthesis control. As sensory feedback is an equally important, integral part of prosthetic research, we will also focus on including tactile force sensors in the fingertips to implement a certain level of adaptivity, and possibly provide tactile feedback to the operator via sensory nerve stimulation.

Aside from our primary objectives for applications in the field of dexterous artificial limb development and industrial telemanipulation, our final, anatomically focused and biomechanically validated results could lay the foundations for advanced prosthetic and biomedical research, most notably by serving as a tool in the development of advanced neural interfaces to identify detailed muscle activity or process motor signals for prosthetic control, and to study muscle activity patterns of the hand during different precision movements towards their reproduction on an algorithmic level.

Acknowledgments

This research is partially supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKPO-16-2017-00002). Development took place in the Robotics Laboratory of the authors' affiliated Faculty.

The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] Shadow Dexterous Hand E1 Series by Shadow Robot Company, 2013, Available: <https://www.shadowrobot.com/products/dexterous-hand/> (last accessed: 29/08/2018)
- [2] Zhang, Ting, et al. "Biomechatronic design and control of an anthropomorphic artificial hand for prosthetic applications." *Robotica* 34.10 (2016): 2291-2308.
- [3] Xu, Zhe, and Emanuel Todorov. "Design of a highly biomimetic anthropomorphic robotic hand towards artificial limb regeneration." *Robotics and Automation (ICRA)*, 2016 IEEE International Conference on. IEEE, 2016.
- [4] Ibrahim A. Kapandji. *The Physiology of the Joints, Volume One, Upper Limb*, 5th ed., Churchill Livingstone, Edinburgh, London, Melbourne and New York, 1982.
- [5] Faudzi, Ahmad Athif Mohd, et al. "Index Finger of a Human-Like Robotic Hand Using Thin Soft Muscles." *IEEE Robotics and Automation Letters* 3.1 (2018): 92-99.
- [6] BodyParts3D site of Anatomography, property of DBCLS, Available: <http://lifesciencedb.jp/bp3d/?lng=en> (last accessed: 12/09/2018)
- [7] Clavero, Juan A., et al. "Extensor mechanism of the fingers: MR imaging–anatomic correlation." *Radiographics* 23.3 (2003): 593-611.
- [8] Backhouse, K. M., and W. T. Catton. "An experimental study of the functions of the lumbrical muscles in the human hand." *Journal of anatomy* 88.Pt 2 (1954): 133.
- [9] Cardoso, Fabiano Nassar, et al. "Imaging the ligaments of the trapeziometacarpal joint: MRI compared with MR arthrography in cadaveric specimens." *American Journal of Roentgenology* 192.1 (2009): W13-W1

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Chemical reaction models of gene expression boundary dynamics in synthetic bacterial colonies

Gregory Szep^{1*}, Paul Grant², Neil Dalchau², Jacob Halatek², Attila Csikasz-Nagy¹ and Andrew Phillips²

¹King's College London

²Microsoft Research Cambridge

*email: gregory.szep@kcl.ac.uk

Keywords: boundary dynamics, orthogonal signalling, synthetic bacteria, chemical reaction model

Robust spatial organisation of cell differentiation during the development of an organism is key to maintaining a fixed number of digits, spinal-discs and segments of various organs [1]. One mechanism by which this organisation may be maintained is by the formation of sharp gene expression boundaries across cell populations in the presence of diffusing signalling molecules that induce switching within cells.

In an effort to understand and control the dynamics and sharpness of such boundaries we recapitulate the context of a multicellular organism as growing colonies of synthetic bacteria. These bacteria have been engineered to express orthogonal signalling pathways that respond to two different inter-cellular signalling molecules [2].

Quantitative measurements were used to build and parametrize a chemical reaction model of inter-cellular synthetic signal transduction. The local concentrations of signalling molecules determine whether the system is bistable or monostable. Separating the model into purely reactive and purely diffusive chemicals allows for the prediction of boundary sharpness and velocity from the flux of signalling molecules. The location of the equilibrium concentration of the diffusive chemicals in state-space is sufficient to predict whether a boundary will be stationary.

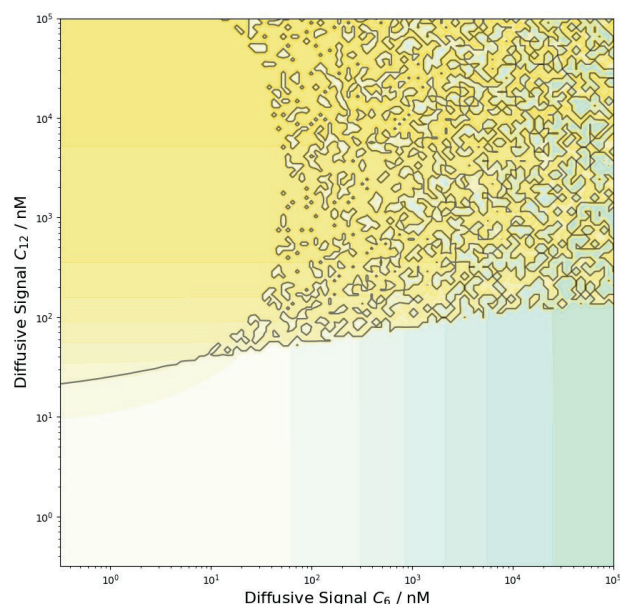


Figure 1 : Contour plot of steady state gene expression levels in cyan/yellow as a function of diffusing signalling molecules. Monostable regions shown in solid colours and bistable regions shown as noisy mixture of yellow and cyan.

An individual cell can be represented as a point in the contour plot shown in Figure 1. It reveals that for sufficiently high local concentrations of both signalling molecules the cell is predicted to be bistable in its gene expression levels. The hysteresis in the bistable region ensures that while the expression of one gene is high, the other is low and vice versa, and that small variations in the concentrations of signalling molecules does not induce switching, making this circuit robust. Switching may be induced by changing the local concentrations of signalling molecules such that the cell crosses the boundary from the bistable to the monostable region. Populations of cells in the presence of diffusing signalling molecules can be represented as a point cloud on top of Figure 1. The uniform equilibrium concentrations of the signalling molecules acts as an attractor for the point cloud; the dynamics are determined by the diffusion equation. The flux of points across the boundary between bistable and monostable regions determines the rate of switching in the population. In the case of countervailing signal gradients this rate also determines the velocity of the gene expression boundary. The desire to calculate fluxes across bifurcation boundaries motivates a formulation of the reaction-diffusion equations in state-space. For problems involving purely diffusive or reactive terms this is relatively straightforward.

Acknowledgements

Gregory Szep is supported by the EPSRC Centre for Doctoral Training in Cross-Disciplinary Approaches to Non-Equilibrium Systems (EP/L015854/1)

References

- [1] Green, J. B. A. and Sharpe, J. 2015. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development*, 142(7), 1203–1211.
- [2] Grant, P. K. et al. 2016. Orthogonal intercellular signaling for programmed spatial behavior. *Molecular Systems Biology*, 12(1), 849–849

BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS

Selected talks

Molecular Filters for Noise Reduction

Luca Laurenti², Attila Csikasz-Nagy^{3,4}, Marta Kwiatkowska², and Luca Cardelli^{1,2}

¹Microsoft Research Cambridge, UK

²Department of Computer Science, University of Oxford, UK

³King's College London

⁴Pazmany Peter Catholic University

Extended Abstract

Biochemical processes such as gene expression are inherently stochastic and must control noise, which presents itself as stochastic fluctuations. These fluctuations can be extrinsic, arising from interactions occurring with other processes in the environment, or intrinsic, resulting from the random timing of the reactions themselves. Molecular processes transform noisy input signals from the environment into output signals through a number of stages, with signals represented by chemical species and each stage implemented by a molecular reaction network. Since each stage can compound the noise, in order to obtain a reliable natural output natural systems must integrate mechanisms that, directly or indirectly, reduce noise or otherwise control it. Examples include signaling cascades, which have been shown to reduce extrinsic fluctuations [4].

Inspired by the concept of low-pass filters in electronics, in this talk, we introduce the concept of molecular filter as a Chemical Reaction Network (CRN) that can reduce the noise of an input signal while still maintaining some of the features of its time evolution, and study noise reduction capabilities of molecular filters in a stochastic setting. We propose three fundamental filter modules and their implementation as stochastic chemical reaction networks. We account for intrinsic as well as extrinsic noise, and derive principles holding for filters when embedded in a general biochemical network, which may include multiple feed-forward and feedback loops. First, we consider linear filters (Figure 1A), implemented by means of at most uni-molecular reactions, and we show that, in case of positive correlation between the elements of the network, they are limited by Poisson levels, that is, the variance of the output signal is lower bounded by its mean. We then show how the presence of feedback loops may improve the performance of linear filters below Poisson levels. Specific models of linear filters have already been studied in the context of closed models of gene expression [3]. Here, while still providing exact analysis based on the solution of the Chemical Master Equation, we generalize the analysis in the referenced papers taking also into account the more general and common case where the dynamics of the other components of the networks are left unspecified, and possibly involving non-linearity and feedback loops. The limitations in the noise suppression capabilities of linear filters motivate us to consider non-linear filters. We introduce a non-linear filter mechanism given by the co-expression of two species that then bind together, called the annihilation module (Figure 1B), which we demonstrate is able to reduce the noise to below Poisson levels. We then propose the annihilation filter (Figure 1C), which combines the properties of the linear filter and annihilation module. We show how the annihilation filter in particular can greatly reduce molecular noise. We observe that co-expression and non-linear degradation are key requirements for such noise reduction. This is important to be stressed as theoretical analysis is usually restricted to linear degradation [2].

Using analytical, numerical and stochastic simulation techniques we demonstrate how the different filters improve robustness of the systems in which they are embedded.

Finally, we discuss that the molecular lters we derive are prevalent in gene expression. For instance, the linear lter, implemented by simply producing and degrading a species at a slow enough rate is a low-pass lter mechanism widely deployed in gene expression to increase robustness, both at the transcription and translation level. Moreover, we find that the annihilation module and annihilation lter are sound models of microRNA regulated gene expression, in the case of correlated expression of microRNAs with the target gene. This supports the hypothesis that microRNA may play a role in increasing robustness and precision of gene expression.

We stress that the focus of this talk is not to offer new models of particular molecular processes, but to identify fundamental and general mechanisms that, at the molecular level, can reduce noise, and to understand their properties and limitations. Thus, our results provide new insight into how biochemical networks control noise. The results of this talk have been published in [1].

References

1. L. Laurenti, A. Csikasz-Nagy, M. Kwiatkowska, and L. Cardelli. Molecular lters for noise reduction. *Biophysical Journal*, 2018.
2. I. Lestas, G. Vinnicombe, and J. Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467(7312):174{178, 2010.
3. B. Munsky, G. Neuert, and A. Van Oudenaarden. Using gene expression noise to understand gene regulation. *Science*, 336(6078):183{187, 2012.
4. M. Thattai and A. van Oudenaarden. Attenuation of noise in ultrasensitive signaling cascades. *Biophysical journal*, 82(6):2943{2950, 2002.

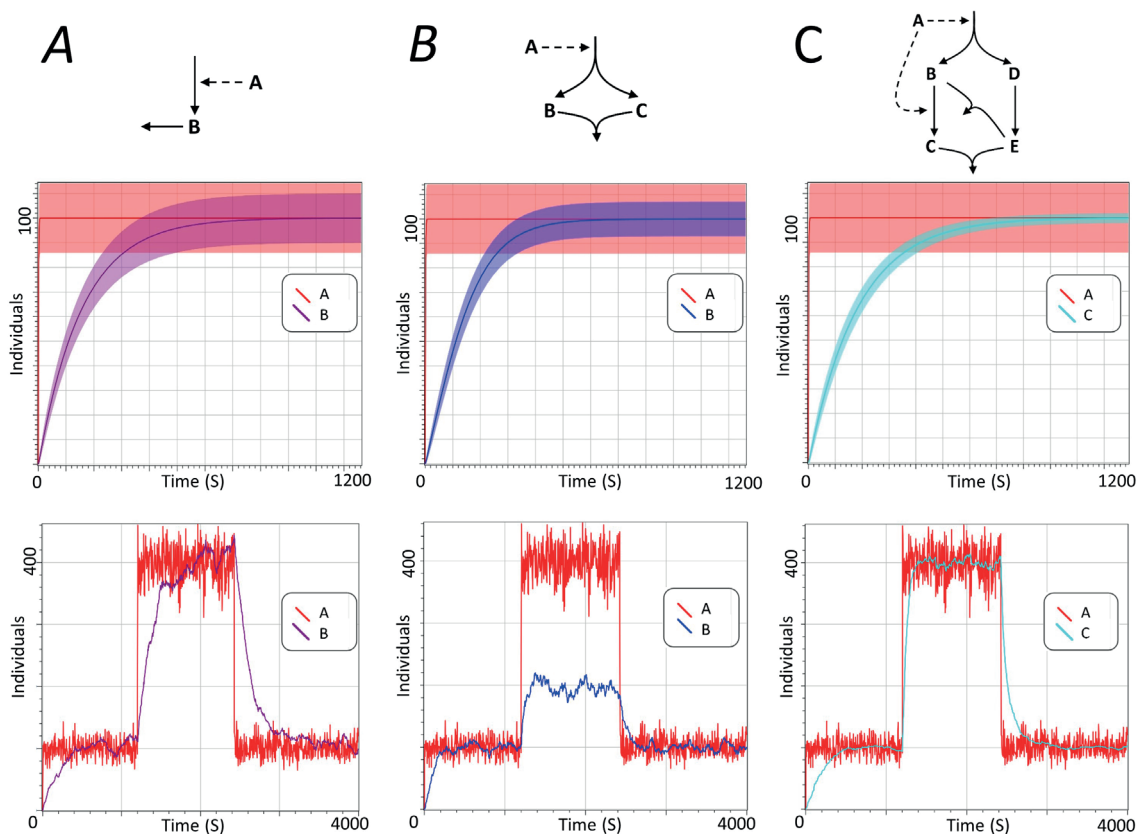


Fig. 1: Comparison of the actions of linear lter, M annihilation module, and annihilation lter on noisy input A generated by reactions $!50 A+A; A \rightarrow !1$. (Top row) Schematic representation of linear lter (A), annihilation module (B), and annihilation lter (C). (Middle row) Plots of expectation and standard deviation of input and output of each lter until time 1200, as estimated by the Linear Noise Approximation (LNA). (Bottom row) Plots of the effect of each lter obtained from a single stochastic simulation until time 4000. (A) The linear lter introduces a delay to buffer stochastic fluctuations and reduce the noise, but only to Poisson levels. (B) The annihilation module improves noise reduction performance but cannot proportionally follow changes in the input. (C) The annihilation lter not only improves the noise reduction capabilities compared to the other modules, but also proportionally follows changes in the input.

Contents of the works presented on the

PROTEIN NETWORK ANALYSIS

conference

PROTEIN NETWORK ANALYSIS

Keynote lecture

Morphisms of chemical reaction networks

Luca Cardelli

Department of Computer Science, University of Oxford, Oxford, UK

We study morphisms (mappings) between chemical reaction networks that establish structural and functional connections between chemical systems. Such morphisms imply kinetic similarity, and yet their properties can be checked statically on the structure of the networks. In particular we can determine structurally that a complex network emulates a simpler network: that is, that it can reproduce the kinetics of the simpler network for all choices of reaction rates and initial conditions of the simpler network. We use this property to relate the kinetics of many common biological networks of different sizes, also relating them to fundamental population algorithms from computing. The emulation property can be used to map out paths in network space that maintain functionality while increasing complexity, thus highlighting possible evolutionary transitions. It can also be used as a baseline to investigate the effects of noise on systems of different complexity.

<http://lucacardelli.name/>

PROTEIN NETWORK ANALYSIS

Keynote lecture

Network analysis and visualization with STRING and Cytoscape

Lars Juhl Jensen

Novo Nordisk Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

Methodological advances have in recent years given us unprecedented information on the molecular details of living cells. However, it remains a challenge to collect all the available data on individual genes and to integrate it with what is described in the scientific literature. The STRING database aims to address this by consolidating known and predicted protein–protein association data for a large number of organisms.

In my presentation, I will give an overview of the STRING database and describe the general approach we use to unify heterogeneous data, provide comparable quality scores for all evidence types, and automatically mine associations from the biomedical literature. I will also show how STRING and associated resources can be used through Cytoscape to visualize your own omics data on a network or to construct and compare, for example, disease-related protein networks. Finally, I will cover some of the new statistical network analysis methods introduced in the latest version of STRING.

<https://jensenlab.org/people/larsjuhljensen/>

PROTEIN NETWORK ANALYSIS

Invited lecture

From protein-protein to human-machine interactions

Jörg Menche

CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Lazarettgasse 14 AKH BT25.3, 1090, Vienna, Austria

From protein interactions to signal transduction, from metabolism to the nervous system: Virtually all processes in health and disease rely on the careful orchestration of a large number of diverse individual components ranging from molecules to cells and entire organs. Networks provide a powerful framework for describing and understanding these complex systems in a wholistic fashion. They offer a unique combination of a highly intuitive, qualitative description, and a plethora of analytical, quantitative tools. In my presentation, I will introduce three ongoing projects of my group, each highlighting a different aspect of how network science can help us understand the pathobiological processes of human disease: First, I will discuss how drug-drug interaction networks can be used to identify basic principles of the cellular response to multiple perturbations. Second, I will show how the mathematical theory of persistent homology reveals a new class of biological networks characterized by a ring-like structure. Lastly, I will present our vision of a virtual reality platform for the next generation of network-based data integration and exploration.

<https://sites.google.com/view/menchelab/home>

PROTEIN NETWORK ANALYSIS

Invited Lecture

Mining the Integrated Connectedness of Biomedical Systems

Natasa Przulj

Department of Computer Science, University College London, WC1E 6BT, London, UK

We are faced with a flood of molecular and clinical data. Various bio-molecules interact in a cell to perform biological function, forming large, complex systems. Large-scale patient-specific omics datasets are increasingly becoming available, providing heterogeneous, but complementary information about cells, tissues and diseases. The challenge is how to mine these interacting, complex, complementary data systems to answer fundamental biological and medical questions. Dealing with them is nontrivial, because many questions we ask to answer from them fall into the category of computationally intractable problems, necessitating the development of heuristic methods for finding approximate solutions.

We develop methods for extracting new biomedical knowledge from the wiring patterns of systems-level, heterogeneous, networked biomedical data. Our methods link the patterns in molecular networks and the multi-scale network organization with biological function. In this way, we translate the information hidden in the wiring patterns into domain-specific knowledge. In addition, we introduce a versatile data fusion (integration) framework that can effectively integrate the information obtained from mining molecular networks with patient-specific somatic mutation data and drug chemical data to address key challenges in precision medicine: better stratification of patients, prediction of driver genes in cancer, and re-purposing of approved drugs to particular patients and patient groups. Our new methods stem from novel network science approaches coupled with graph-regularized non-negative matrix tri-factorization, a machine learning technique for dimensionality reduction and co-clustering of heterogeneous datasets. We utilize our new framework to develop methodologies for performing other related tasks, including disease re-classification from modern, heterogeneous molecular level data, inferring new Gene Ontology relationships, and aligning multiple molecular networks.

<http://www0.cs.ucl.ac.uk/staff/natasa/index.html>

PROTEIN NETWORK ANALYSIS

Invited lecture

Answering questions in biology and medicine by making inferences on networks

Alberto Paccanaro

Centre for Systems and Synthetic Biology, Department of Computer Science, Royal Holloway, University of London, Egham, TW20 0EX, UK

An important idea that has emerged recently is that a cell can be viewed as a set of complex networks of interacting bio-molecules and genetic disease is the result of abnormal interactions within these networks. In this talk, I'll present novel computational methods for answering questions in systems biology and medicine which can all be phrased in terms of inference and structure discovery in such large scale networks. These methods are based and extend recent developments in the areas of machine learning (particularly semi-supervised learning and matrix factorization), graph theory and network science. I'll show how these computational techniques can provide effective solutions for: 1) quantifying similarity between heritable diseases at molecular level using exclusively disease phenotype information; 2) disease gene prediction; 3) drug side-effect prediction.

<https://www.cs.rhul.ac.uk/home/alberto/>

PROTEIN NETWORK ANALYSIS

Selected talks

MiStlmm: an agent-based simulation tool to study the self-nonself discrimination of the adaptive immune response

Csaba Kerepesi^{1,4,*}, Tibor Bakács², Tamás Szabados³

¹Institute for Computer Science and Control, Hungarian Academy of Sciences, Budapest, Hungary

²Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

³Department of Mathematics, Budapest University of Technology and Economics, Budapest, Hungary

*Email: kerepesi@sztaki.hu

Keywords: immune system simulation; self-nonself discrimination, self-centered model, agent-based model

Background: We are witnessing a major change in immunology's conceptual character from an emphasis on immunity as a defense to immunity as an interface-of-exchange. Immunity should be regarded as a communicative system of the internal homeostasis which perceives and then mediates environmental information (organic and inorganic; internal and external) [1]. To handle this complexity there is an increasing need for complex computational models to perform in silico experiments as an adjunct to in vitro and in vivo experiments. One of the key points of immunity is the concept of self-nonself discrimination. We proposed first that in order to recognize self and non-self, T lymphocytes should recognize the much smaller set of self antigens, rather than the practically unlimited non-self antigen universe [2, 3]. The immune system is continuously in a state of delicate balance between tolerating self and attacking non-self. If this balance is perturbed, autoimmune reactions occur. Immunological tolerance is rooted in regulatory immune cell subsets, suppressive cytokines, and immune checkpoint pathways [4]. A good example for the delicate balance of immune tolerance and intolerance – and for the importance of this research area – is the ambiguous results of The Cancer Immunotherapy Revolution [5], in which the newly approved immunotherapies manipulate components of the immune system to attack tumors. Hundreds of clinical trials are underway to improve responses and success stories of terminal cancer patients defying the odds and achieving complete remissions are accumulating. Unfortunately, the manipulation of the immune system has also resulted in a major safety issue: the iatrogenic immunerelated adverse events (irAEs). As a result of the impaired self-tolerance, irAEs may present with a broad clinical spectrum that mainly involves the gut, skin, endocrine glands, liver, and lung but can potentially affect any tissue, and their incidence may reach up to 90% of patients [6, 7]. In order to aid in the qualitative characterization and examination of the delicate immune balance, we have developed MiStlmm computer program, which is capable to simulate the complex processes of self-nonself discrimination of the adaptive immune system. We know that a computer model can not reliably simulate the whole immune system, however, simulating areas of interest can be useful testing ideas to help in the design of in vivo and in vitro experiments [8].

Implementation of MiStlmm: MiStlmm uses agent-based modeling technique [9] and it can simulate some aspects of humoral immune response along with its major components, including T cells, B cells, antibodies, danger signals, interleukins, self cells and foreign antigens. These simulation components (called “agents”) determine the nodes of a dynamic immune network where links are the potential interactions between two elements. The immune network changes step by step (in time) driven by random events. Using the terminology of [10], a model simulated by MiStlmm is an agent-based model that is in part “individual particle based-stochastic”, and in part “particle number stochastic”. A great advantage of such a model is that it can easily incorporate the most important types of cells and molecules together with their essential features and simulation

events that play important roles in immune reactions. In such a simulation events – for example interactions of components – occur at random. A stochastic model fits well with the affinity maturation of B lymphocytes in which random events are perhaps the most characteristic. It is also suitable to model the development of the regulatory T cell population and the random selection of specific T cell clones. To simplify things, we chose the humoral adaptive immune system since the humoral phase (blood or lymph) may be considered spatially homogeneous; thus a microscopic spatial volume may represent the whole phase well. A major advantage of this approach is that it is not necessary to describe the actual spatial positions and spatial motions in the model. Instead, model components (agents) randomly choose one of the other components as interaction partners, because any components are close enough to become engaged in an interaction.

First application of MiStlmm: As the first application of MiStlmm, we have simulated two different immune models and then we have compared performances of them in the mean efficacy of self-nonself discrimination. The first model is called nonself centered or Conventional Role of Self (CRS) where even a primary immune reaction depends on the recognition of non-self antigens by T and B cell receptors [11, 12, 13]. The role of self in this model is that the great majority of autoreactive T and B cell clones are selected and purged from the immune system [14]. The second model called self-centered or Enhanced Role of Self (ERS) which is based on our previously published “one-signal model” [3]. We proposed that model (hypothesis) when we have been seeking the answer for three unresolved paradoxes of immunology:

- (Q1) How can a tiny fraction of human genome effectively compete with a vastly larger pool of mutating pathogen DNA [15]?
- (Q2) Considering the fact, that average 3 mutations are formed each of the 10^{16} times the cell's 3×10^9 DNA base pairs are duplicated during a human lifetime [16], “why does cancer occur so infrequently”?
- (Q3) Considering the facts that T cells require three to five days to attain fighting strength (because they are rare, short-lived, and their doubling time is at least 6 h [17]), yet how can a T cell response be measurable in the lymph nodes draining the infection site within 12 to 18 h [17].

In order to explain these paradoxes, we have suggested a new T cell model [3] that we can summarize in the following. We have postulated that a dynamic steady state, a so-called coupled system is formed through low affinity complementary TCR–MHC interactions between T cells and host cells. Under such condition, it is sufficient to recognize what is self in order to attack nonself (answer to Q1). We have postulated that the evolutionary pressure driving the creation of the T cell receptor (TCR) repertoire was primarily the homeostatic surveillance of the genome (answer to Q2). The new model implies that a significant fraction of the naive polyclonal T cells is recruited into the first line of defense from the very outset of an infection (answer for Q3). The computational variant of our hypothesized T-cell model is the ERS model, presented in this paper. Main goal of the simulation experiments of the current paper is showing that the ERS model matches real patterns and additionally to analyze how the two models (CRS and ERS) cope against a critical primary infection. There are several available computational immune system simulation tools that are capable to simulate a conventional or standard immune model (like the CRS model), but we have not found one that was designed for simulating the ERS model and compare it with a conventional model. That was the main reason why we have developed MiStlmm. Simulation results by MiStlmm have shown that the ERS model does not develop autoimmune reactions despite the existence of the hypothesized TCR–MHC interaction between T cells and self antigens in the model. Our results also have shown that the ERS model gives better results to overcome a critical primary infection, answering the paradox “how can a tiny fraction of human genome effectively compete with a vastly larger pool of mutating pathogen DNA?”

Conclusion: We hope that our results will encourage investigations to make in vitro and in vivo experiments clarifying questions about self-nonself discrimination of the adaptive immune system. We also hope that MiStlmm or some concept in it will be useful for implementation and/or comparing other immune models by other researchers.

Acknowledgments

We are grateful to Dr. Andrea Ladányi for her useful criticism and to Prof. Péter Csermely for his useful advice. We thank former students Gábor Szabados, Tamás Kiss, Kristóf Hörömpöly, and Endre Szecsei for their contributions to developing the software. CK is supported by the „2018-1.2.1-NKP-00008: Exploring the Mathematical Foundations of Artificial Intelligence” grant of Ministry of Human Capacities in Hungary.

References

1. Tauber, A.I.: Immunity: the Evolution of an Idea. Oxford University Press, Oxford (2017)
2. Bakács, T., Mehrishi, J.N., Szabados, T., Varga, L., Szabó, M., Tusnády, G.: T cells survey the stability of the self: a testable hypothesis on the homeostatic role of TCR-MHC interactions. *Int Arch Allergy and Immunol* 144, 171–182 (2007)
3. Szabados, T., Bakács, T.: Sufficient to recognize self to attack non-self: blueprint for a one-signal T cell model. *J Biol Syst* 19, 299–317 (2011)
4. Topalian, S.L.: Targeting immune checkpoints in cancer therapy. *Jama* 318(17), 1647–1648 (2017)
5. Kelly, P.N.: The Cancer Immunotherapy Revolution. American Association for the Advancement of Science (2018)
6. Puzanov, I., Diab, A., Abdallah, K., Bingham, C., Brogdon, C., Dadu, R., Hamad, L., Kim, S., Lacouture, M., LeBoeuf, N., et al.: Managing toxicities associated with immune checkpoint inhibitors: consensus recommendations from the Society for Immunotherapy of Cancer (SITC) Toxicity Management Working Group. *Journal for immunotherapy of cancer* 5(1), 95 (2017)
7. Brahmer, J.R., Lacchetti, C., Thompson, J.A.: Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: American Society of Clinical Oncology Clinical Practice Guideline Summary. *Journal of oncology practice* 14(4), 247–249 (2018)
8. Celada, F., Seiden, P.E.: A computer model of cellular interactions in the immune system. *Immunol Today* 13, 56–62 (1992)
9. Bonabeau, E.: Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences* 99(suppl 3), 7280–7287 (2002)
10. Germain, R.N., Meier-Schellersheim, M., Nita-Lazar, A., Fraser, I.D.C.: Systems biology in immunology – A computational modeling perspective. *Annu Rev Immunol* 29, 527–585 (2011)
11. Bretscher, P., Cohn, M.: A theory of self-nonself discrimination: Paralysis and induction involve the recognition of one and two determinants on an antigen, respectively. *Science* 169(3950), 1042–1049 (1970)
12. Janeway, C., Murphy, K.P., Travers, P., Walport, M.: Immunobiology. Garland Science (2008)
13. Paul, W.E.: Fundamental Immunology. Lippincott Williams & Wilkins (2008)
14. Faro, J., Velasco, S., González-Fernández, Á., Bandeira, A.: The impact of thymic antigen diversity on the size
15. Langman, R.E.: Molecular economy and antibody function: the evolution of a Protecton. Springer (1992)
16. Bredberg, A.: Cancer resistance and peto’s paradox. *Proc Natl Acad Sci USA* 106, 51 (2009)
17. AK, A.K.A., Lichtman, A.H.: Basic Immunology: Functions and Disorders of the Immune System, 2nd edn. Saunders, Philadelphia (2004)

PROTEIN NETWORK ANALYSIS

Selected talks

Integrating molecular databases in a proteome-wide simulation environment.

Bence Keömley-Horváth¹, Simone Rizzeto², István Reguly¹, and Attila Csikász-Nagy^{1,3}

¹Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, Budapest, Hungary

²Viral Immunology Systems Program, Kirby Institute School of Medical Sciences, UNSW, Sydney, Australia

³Randall Division of Cell and Molecular Biophysics, King's College London, London, UK

⁴3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

In the last few decades due to new high-throughput techniques, such as mass spectrometry, we gained a large amount of data about Protein-Protein Interactions (PPI) and Domain-Domain Interactions (DDI). This opened the opportunity for computational methods to predict protein complexes. Clustering methods produce only qualitative information about the complexes even though they are formed dynamically in various parts of the cells. The number of complexes formed is limited by proteins which have many possible binding partners while present in low abundance.

Here we propose a new version of SiComPre¹, a simulation based protein complex prediction method, that is based on Gillespie's multiparticle algorithm. This combines the classic Gillespie's algorithm with diffusion by splitting the simulation space into sub-volumes (SVs) which are small enough that we can assume that the reactant are able to find each others. In given times the molecules diffuse between the SVs^{2,3}.

The previous SiComPre program has some limitations; the model of the cell is only two dimensional and it could utilize only one GPU, limiting the abundance of the simulateable proteins, because of the running time. We want to be able to utilize multiple GPUs, if this is necessary. In the new version, we want to implement a three dimensional cell model, thus the diffusion will be much less limited. The new version is developed in a more Object-Oriented Programming way, which makes the code more flexible for the future features, such as more specified internal bounds in complexes, protein level conformation changes, complex specific diffusion properties.

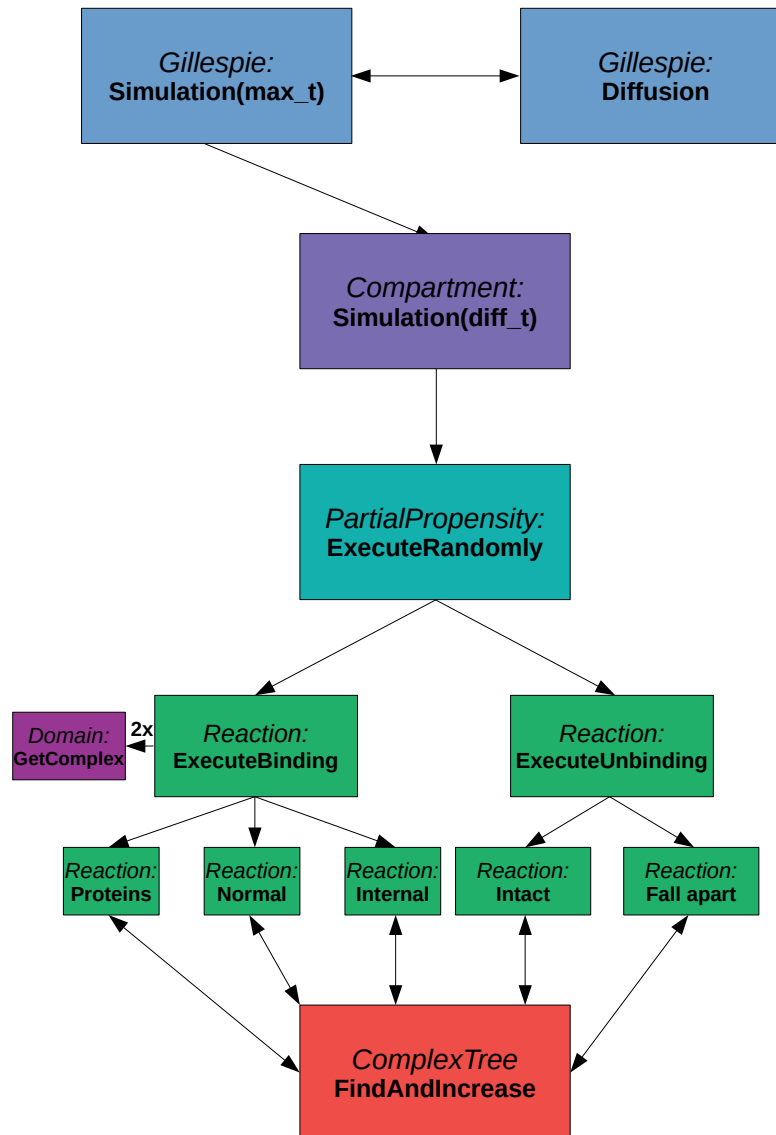
Furthermore, a proteome-wide diffusion coefficient database is not available for any organisms, so we developed a method that predicts these constants from protein masses with the help of Stokes-Einstein equation. Here we assume that the proteins are globular, this is accurate enough to know the scales⁴.

Algorithm

We use the statistically more efficient Partial Propensity method⁵ instead of the original Gillespie algorithm⁶. The advantage of this algorithm is that it makes the search for the next reaction proportional to the number of the reactants (protein domains) instead of reactions. The cost of this method is that the searching and refreshing are more complicated from the implementation point of view.

The central class for the entire algorithm is the Gillespie class, which stores every necessary data for the algorithm in specific classes, the compartments, SVs, reactions, proteins, domains and the important parameters like the time length of the simulation the time density of the diffusion. The complexes are stored in

a binary search tree, where the key is an integer vector which represents the binding order and indirectly the composition of the complex. The Complex class has several attributes, the composition, bindings, free binding sites, abundance, diffusion propensities and the key. The proteins are also represented as a Complex class instance, but the binding array is empty and the composition is the self identification. We are in the final steps in the development of the algorithm, the algorithm is now working but it is unstable at the moment. The steps



of the algorithm are in the 1. Figure:

Figure 1. The entire simulation is launched from the simulation member function of the Gillespie class. It calls every compartments to execute its simulation function for a diffusion time step, when these are finished it calls the diffusion member function. Inside the Compartment's simulation function, the selection of the next reaction happens in two steps, according to the Partial Propensity, first we chose the reaction group (instance of PartialPropensity class) and then we select the exact reaction. Here we decide if the binding or the unbinding reaction will be executed. For the binding we need to get two complexes from the reactant domains randomly, if we get two proteins then the protein binding function will be executed, if at least on of them is complex then the normal binding will be executed and the last possibility is the internal binding. In the unbinding case there are two possibilities, the complex can fall apart into two complexes or can remain intact but loses one internal bound. After every reaction we search for the output complex(s) and we increase the abundance or we insert it/them to the tree. If we insert them back we have to finish the creation of every attribute.

Acknowledgement:

The research has been supported by the European Union, co-financed by the European Social Fund(EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

1. Rizzetto S, Priami C, Csik'asz-Nagy A. Qualitative and Quantitative Protein Complex Prediction Through Proteome-Wide Simulations. PLoS Comput Biol 11(10): e1004424. doi:10.1371/journal.pcbi.1004424, 2015
2. Stundzia, A.B. and Lumsden, C.J. Stochastic simulation of coupled reaction-diffusion processes. Journal of Computational Physics (1996), 127, 196-207.
3. Simone Rizzetto, Petros Moyseos, Bianca Baldacci, Corrado Priami, Attila Csik'asz-Nagy Contextdependent prediction of protein complexes by SiComPre npj Systems Biology and Applications volume 4, Article number: 37 (2018)
4. M. E. Young P. A. Carroad R. L. Bell Estimation of diffusion coefficients of proteins Biotechnology and Bioengineering, Vol. XXII, Pp. 947-955 (1980)
5. Rajesh Ramaswamy, N'elido Gonz'alez-Segredo, Ivo F. Sbalzarini A new class of highly efficient exact stochastic simulation algorithms for chemical reaction networks arXiv:0906.1992 [q-bio.QM]
6. Daniel T. Gillespie A general method for numerically simulating the stochastic time evolution of coupled chemical reactions Journal of Computational Physics (1976), 22, 403-434

PROTEIN NETWORK ANALYSIS

Selected talks

Exploring inhibitory logics in protein-protein interaction networks

Fridolin Gross and Andrea Ciliberto
IFOM, Via Adamello 16, 20139, Milan, Italy

We explore models of molecular networks, where activation and inhibition take place by direct binding. Our analysis starts from a specific example, the mitotic checkpoint (also called spindle assembly checkpoint), which is a signaling pathway that ensures faithful chromosome segregation. Mitotic checkpoint proteins inhibit the anaphase-promoting complex (APC/C) and its activator Cdc20 to prevent precocious anaphase. Checkpoint signaling leads to a complex of APC/C, Cdc20, and checkpoint proteins, in which the APC/C is inactive. In principle, this final product of the mitotic checkpoint can be obtained via different pathways, whose relevance still needs to be fully ascertained experimentally. Here, we use mathematical models to compare the implications on checkpoint response of the possible pathways leading to APC/C inhibition. We identify a previously unrecognized funneling effect for Cdc20, which favors Cdc20 incorporation into the inhibitory complex and therefore promotes checkpoint activity. Furthermore, we find that the presence or absence of one specific assembly reaction determines whether the checkpoint remains functional at elevated levels of Cdc20, which can occur in cancer cells. Finally, we extend our results to generic protein-protein interaction networks.

PROTEIN NETWORK ANALYSIS

Selected abstracts

Tumor growth modeling using formal reaction kinetics analogy

Dániel András Drexler^{1*}

¹Physiological Controls Research Center, University Research, Innovation and Service Center, Óbuda University

*E-mail: drexler.daniel@nik.uni-obuda.hu

Keywords: tumor model, reaction steps, washout, drug resistance

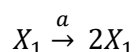
Application of certain areas of mathematics and control theory is becoming an essential part of modern medicine. Dynamical models (i.e., differential equations) of diseases and physiological phenomena can be used for prognosis, simulation of the effect of drugs, and even therapy design. Moreover, physiological modeling is a fundamental component of artificial organ design.

Tumor growth modeling may have great impact on antitumor therapy research. The tumor growth models can be used to simulate the effect of a drug; through parametric identification, the model parameters can be used to classify drug and tumor types, while appropriately simple, but still efficient tumor growth models can be used to design effective therapies using the results of control theory. Creating a relatively simple tumor growth model with good modeling power thus may lead to more optimized tumor treatment therapies.

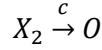
The most used tumor growth model which models the effect of antiangiogenic drugs was created in 1999 at the Harvard Medical school by Philip Hahnfeldt and his collaborates [1]. The Hahnfeldt model is created to describe the relationship between the tumor vasculature and tumor volume. It uses Gompertzian growth function to describe the tumor growth, and the vasculature growth dynamics is also described by nonlinear terms. The model has been modified and conditions of effective therapy were derived in [2]. The Hahnfeldt model is extended with time delays and generalized to describe chemotherapy and radiotherapy in [3].

Although the Hahnfeldt model is very popular, and several therapies have been designed using control theory based on this model (see e.g., [4,5]), the model has some drawback. First, the necrotic (dead) tumor volume is not considered in the model. During the diagnosis process, proliferating (living) tumor volume is measured together with necrotic tumor volume, thus necrotic tumor volume dynamics modeling may have great practical importance. The other drawback of the Hahnfeldt model is that it contains many nonlinear terms which complicates controller design. Although necrotic tumor volume modeling may be negligible during chemotherapy due to the washout of the necrotic tumor cells, the washout process is inhibited during antiangiogenic therapy due to the lack of appropriate vasculature.

A simple tumor growth model was created in [6] based on three formal kinetic reactions. The first reaction defines the tumor growth with the formal reaction step

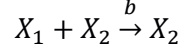


where X_1 is the fictive species that corresponds to the proliferating tumors. Denoting the tumor volume by x_1 , and considering this reaction step with mass-action kinetics, we get the differential equation $\dot{x}_1 = ax_1$. The second reaction defines drug depletion with the formal reaction step



where X_2 is the fictive species that corresponds to the drug level in the patient. Denoting the drug level by x_2 , and considering this reaction step with mass-action kinetics, we get the differential equation $\dot{x}_2 = -cx_2$.

The third reaction step defines the effect of the drug with the formal reaction step

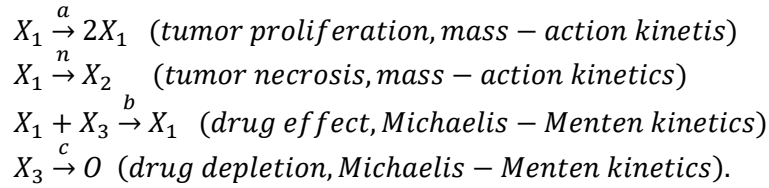


that results in the differential equation $\dot{x}_1 = ax_1 - bx_1x_2$. Considering the drug injection as the time function u which increases the drug level, we get the differential equations of the tumor growth model:

$$\begin{aligned}\dot{x}_1 &= ax_1 - bx_1x_2 \\ \dot{x}_2 &= -cx_2.\end{aligned}$$

This model contains two linear and one bilinear term, thus it is simple enough for controller design (see e.g., [7,8]), and still has appropriate modeling power by modeling the relevant physiological phenomena, as described by the fictional reaction steps. However, this model does not consider the necrotic tumor volume.

An extended version of this tumor model has been published in [9] that models the necrotic tumor volume as well, while considering the effect of the drug and the clearance of the drug using Michaelis-Menten kinetics, which results in a physiologically more accurate model. The fictive species in the extended model are X_1 , X_2 and X_3 , and correspond to the proliferating tumor volume, necrotic tumor volume and drug serum level, respectively. The formal reaction steps that define the model are

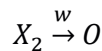


The corresponding differential equations are

$$\begin{aligned}\dot{x}_1 &= (a - n)x_1 - b \frac{x_1x_3}{ED_{50} + x_3} \\ \dot{x}_2 &= nx_1 + b \frac{x_1x_3}{ED_{50} + x_3} \\ \dot{x}_3 &= -c \frac{x_3}{K_B + x_3} - b_\kappa \frac{x_1x_3}{ED_{50} + x_3}\end{aligned}$$

where ED_{50} and K_B are the Michaelis-Menten parameters for the last two reaction steps, respectively, and the time function u is the injection input, while the time functions x_1 , x_2 and x_3 describe the proliferating tumor volume, necrotic tumor volume and drug serum level, respectively. This model describes the necrotic tumor volume, models most of the relevant physiological phenomena, and is still relatively simple. The qualitative analysis of the model with the application of simple feedback control has been carried out in [10], where conditions for the model parameters for the existence of an effective therapy were derived.

This model can be modified to incorporate the effect of necrotic tumor cell washout by adding the formal reaction step



as it was already shown in [10]. By considering the rate b as a state variable that increases over time if there is no inhibitor present, but decreases if there is inhibitor present, we can model the drug resistance dynamics. By adding the necrotic tumor cell washout and resistance dynamics, the model will be able to describe the effect of chemotherapeutic drugs as well (the model was originally created to model the effect of antiangiogenic drugs, in which case these two effects were negligible).

The formal reaction kinetics analogy was found to be useful in the creation of tumor growth models. The other advantage of the analogy is that the formal reaction steps are comprehensible for experts who are not familiar with differential equations (e.g., medical doctors). Moreover, since the dynamical system is based on reaction steps, the positivity property is guaranteed inherently due to the positivity of kinetic systems.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 679681).

References

- [1] P. Hahnfeldt, D. Panigrahy, J. Folkman, and L. Hlatky. Tumor development under angiogenic signaling: A dynamical theory of tumor growth, treatment response, and postvascular dormancy. *Cancer Research*, 1999, 59, 4770–4775.
- [2] A. d’Onofrio and A. Gandolfi. Tumor eradication by antiangiogenic therapy: analysis and extensions of the model by Hahnfeldt et al. (1999). *Mathematical Biosciences*. 2004, 191(2), 159–184.
- [3] P. Amster, L. Berezansky, L. Idels. Stability of Hahnfeldt angiogenesis models with time lags. *Mathematical and Computer Modeling*, 2012, 55(9-10), 2052–2060.
- [4] L. Kovács, A. Szeles, J. Sápi, D. A. Drexler, I. Rudas, I. Harmati, Z. Sápi. Model-based angiogenic inhibition of tumor growth using modern robust control method. *Computer Methods and Programs in Biomedicine*, 2014, 114, e98–e110.
- [5] Sápi, J., Drexler, D. A., Harmati, I., Sápi, Z., and Kovács, L. Qualitative analysis of tumor growth model under antiangiogenic therapy – choosing the effective operating point and design parameters for controller design. *Optimal Control Applications & Methods*, 2016, 37, 848–866.
- [6] D. A. Drexler, J. Sápi, L. Kovács. A minimal model of tumor growth with angiogenic inhibition using bevacizumab. in *Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics*, 2017, 185–190.
- [7] D. A. Drexler, J. Sápi, L. Kovács. Positive control of a minimal model of tumor growth with bevacizumab treatment. In *Proceedings of the 12th IEEE Conference on Industrial Electronics and Applications*, 2017, 2081–2084.
- [8] D. A. Drexler, J. Sápi, L. Kovács. H^∞ control of nonlinear systems with positive input with application to antiangiogenic therapy. in *Proceedings of the Joint 9th IFAC Symposium on Robust Control Design and 2nd IFAC Workshop on Linear Parameter Varying Systems*. 2018, 219–224.
- [9] D. A. Drexler, J. Sápi, L. Kovács. Modeling of tumor growth incorporating the effects of necrosis and the effect of bevacizumab. *Complexity*, 2017, 1–11.
- [10] D. A. Drexler, I. Nagy, V. Romanovski, J. Tóth, L. Kovács. Qualitative analysis of a closed-loop model of tumor growth control. In *Proceedings of the IEEE 18th International Symposium on Computational Intelligence and Informatics*, 2018.

PROTEIN NETWORK ANALYSIS

Selected talks

A Bayesian Ranking Approach for Drug-Target Interaction Prediction and a Web Application Containing the Predictions

Ladislav Peška^{1*}, Krisztian Buza^{2*}, Júlia Koller^{3*}

¹ Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

² ELTE – Eötvös Loránd Tudományegyetem, Budapest, Hungary

³ Institute of Genomic Medicine and Rare Disorders, Semmelweis University, Budapest, Hungary

*E-mail: peska@ksi.mff.cuni.cz, buza@biointelligence.hu, jkoller@biointelligence.hu

Keywords: drug repositioning, drug-target interactions, machine learning, Bayesian personalized ranking

Background and Objective: In our work [1] we considered in silico prediction of drug-target interactions (DTI), which could provide valuable information and speed-up the process of drug repositioning – finding novel usage for existing drugs. In particular, we focused on machine learning algorithms supporting drugcentric repositioning approach.

Methods: We proposed Bayesian Ranking Prediction of Drug-Target Interactions (BRDTI), a per-drug ranking-based method, which reflects the needs of drug-centric repositioning research better than conventional drug-target prediction approaches. Our method is based on Bayesian Personalized Ranking (BPR) [2] which has been shown to be an excellent approach for various preference learning tasks, however, it has not been used for DTI prediction previously. In order to successfully deal with DTI challenges, we extended BPR by proposing: (i) the incorporation of target bias, (ii) a technique to handle new drugs and (iii) content alignment to take structural similarities of drugs and targets into account.

Results: Evaluation on five benchmark datasets [3], [4] according to 5x10-fold cross-validation shows that BRDTI outperforms several state-of-the-art approaches in terms of per-drug normalized discounted cumulative gain (nDCG) and area under receiver-operator characteristic curve (AUC). BRDTI results w.r.t. nDCG are 0.929, 0.953, 0.948, 0.897 and 0.690 for G-Protein Coupled Receptors (GPCR), Ion Channels (IC), Nuclear Receptors (NR), Enzymes (E) and Kinase (K) datasets respectively. Additionally, BRDTI significantly outperformed other methods (BLM-NII, WNN-GIP, NetLapRLS and CMF) w.r.t. nDCG in 17 out of 20 cases.

Furthermore, according to the schema in Fig. 1, BRDTI was also shown to be able to predict novel drugtarget interactions not contained in the original datasets. The average recall at top-10 predicted targets for each drug was 0.762, 0.560, 1.000 and 0.404 for GPCR, IC, NR, and E datasets respectively.

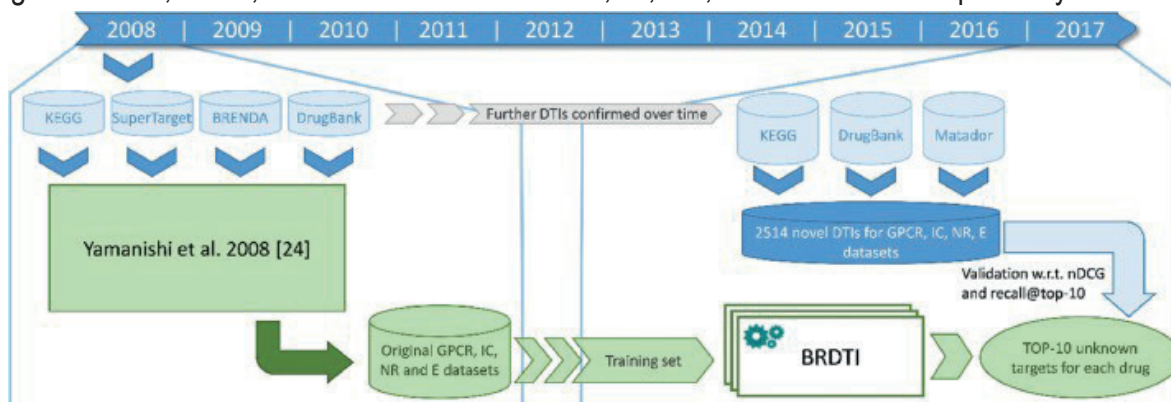


Fig. 1. Evaluation workflow using recently confirmed DTIs.

Moreover, we set up a web page (<http://www.ksi.mff.cuni.cz/~peska/BRDTI/newDTI/>) that contains (i) top-50 predictions for Enzymes, GPCR, IC and NR, respectively; and (ii) top-10 predicted interactions for each drug. In the presentation, we will introduce this web application and point out some of the predicted interactions and discuss them in detail in order to demonstrate the feasibility of predictions.

Conclusions and Outlook: Based on the evaluation, we can conclude that BRDTI is an appropriate choice for researchers looking for an in silico DTI prediction technique to be used in drug-centric repositioning scenarios. BRDTI Software and supplementary materials are available online at www.ksi.mff.cuni.cz/~peska/BRDTI . In our current work, we are investigating other approaches for DTI, such as bipartite local models with hubness-aware regressors [5]. Other machine learning techniques for DTI prediction include approaches for DTI include Bayesian matrix factorization [6] and restricted Boltzmann machines [7].

References

- [1] L. Peška, K. Buza, J. Koller. Drug-target interaction prediction: A Bayesian ranking approach. *Computer Methods and Programs in Biomedicine*, 2017, 152, 15-21
- [2] S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme. BPR: Bayesian personalized ranking from implicit feedback. *Proc. Twenty-Fifth Conf. Uncertain. Artif. Intell.*, 2009, 452-461
- [3] K. Bleakley, Y. Yamanishi. Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics*, 2009, 25.18, 2397-2403
- [4] J.T. Metz, E.F. Johnson, N.B. Soni, P.J. Merta, L. Kifle, P.J. Hajduk. Navigating the kinome. *Nature chemical biology*, 2011, 7.4, 200
- [5] K. Buza, L. Peška. Drug–target interaction prediction with Bipartite Local Models and hubness-aware regression. *Neurocomputing*, 2017, 260, 284-293
- [6] B. Bolgar, P. Antal. Bayesian matrix factorization with non-random missing data using informative gaussian process priors and soft evidences. *J. Mach. Learn. Res.*, 2016, 52, 25-36
- [7] Y. Wang, J. Zeng. Predicting drug-target interactions using restricted boltzmann machines. *Bioinformatics*, 2013, 29(13), i126-i134

PROTEIN NETWORK ANALYSIS

Selected talks

Binding sites in disordered regions might contribute to the complexity of the postsynaptic protein network

Annamária Kiss-Tóth^{1,2*}, Balázs Ligeti¹, Gergely Lukács¹, Zoltán Gáspári¹

¹ Pázmány Péter Catholic University, Faculty of Information Technology and Bionics

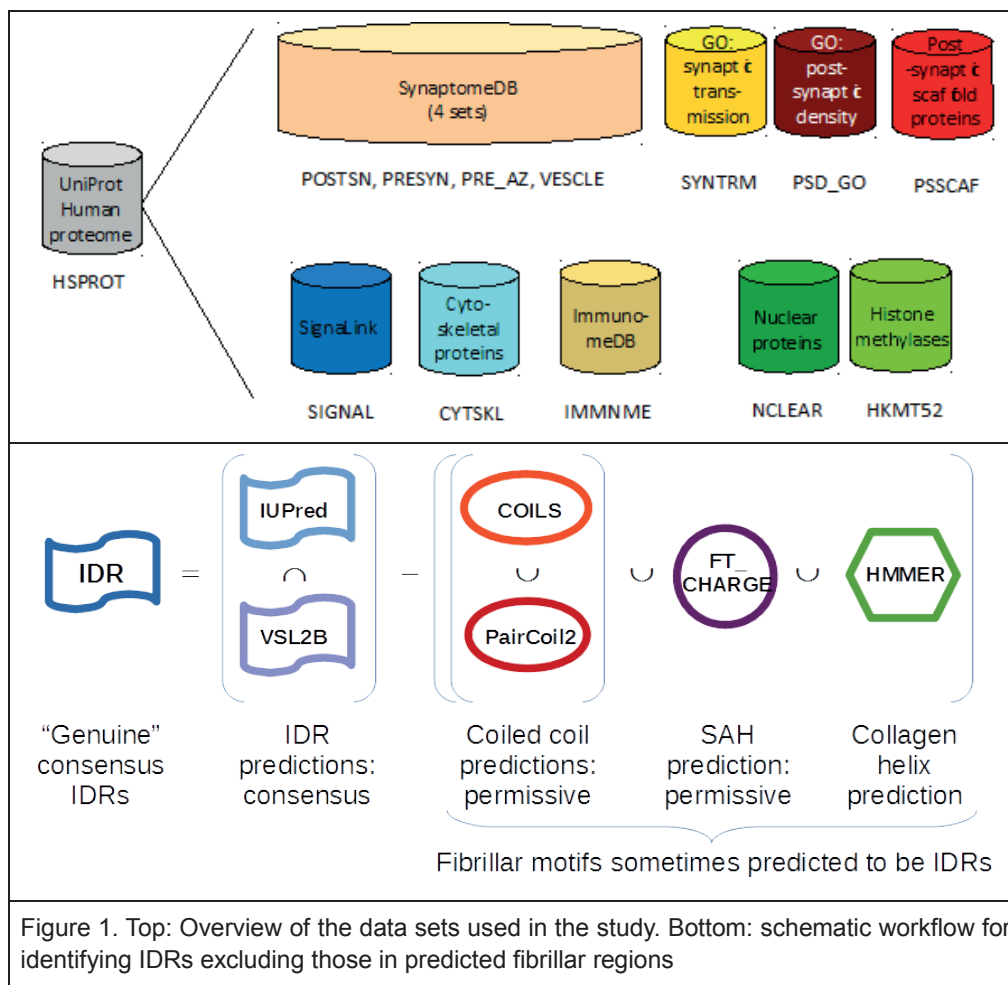
² 3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: kiss-toth.annamaria@itk.ppke.hu

Keywords: postsynaptic density, intrinsic disorder tendency, binding sites, RNA binding proteins

The synapse is a structure, which provides a way for neuronal communication. Postsynaptic density (PSD) is a characteristic part of excitatory chemical synapses. It forms a disk-shaped electrodense entity about 200-800 nm wide and 30-50 nm thick beneath the postsynaptic membrane. This dense, protein-based network provides a complex connection between the intracellular portions of membrane receptors and adhesion molecules and the cytoskeleton (1). The postsynaptic density is responsible for modulating signal transduction and orchestrating a number of processes linked to learning and memory. The postsynaptic density protein network is substantially more than a passive scaffold for receptors and signal transducers, as many observations show that it is undergoing dynamic reorganization during the course of various neural processes. The size of postsynaptic density correlates with the strength of the synapse, and morphological differences between the PSDs are observed in the different brain regions. The PSD can most likely be imagined as a supramolecular machine capable of integrating and transmitting signals via dynamic reorganization. The underlying mechanisms of this most likely include competitive binding events, allostery and cooperativity at the level of proteins and the network might act as an amplifier and integrator of these. Recent experimental observations suggest the possible role of phase transitions in both post- and presynaptic processes (2, 3).

In this project, we investigated whether disorder plays a role in the organizing, variability of postsynaptic density proteins and the dynamic rearrangement of postsynaptic density. While there are domains such as PDZ domains that have been shown to be abundant in postsynaptic proteins, we do not know in the synaptom a detailed analysis of the role of disorder tendency. In order to achieve a more accurate result, we considered the consensus of two different prediction methods to minimize the number of false positives. We chose IUPred and VSL2B, which are considered some of the most reliable disorder prediction algorithms. In addition to this, we set a so- called 'genuine disorder' instead of ordinary disorder. This means that out of the disorderly segments we have filtered motifs that were often predicted to be disordered in the past, such as coiled-coil structures. We believe that this will ensure that we obtain results that are more relevant to the in vivo protein disorder. In order to bring the results into sufficient light, we have chosen a number of external reference sets: the entire human proteome, the immunome, consisting primarily of globular multidomain proteins, nuclear proteins and a special subset of histone methylases, which are rich in known segregated segments and signaling pathways participating proteins (Signalink) and cytoskeletal proteins (Figure 1). The latter two groups were chosen because they are related to the function and organization of synaptom. Binding regions within disordered segments were identified with ANCHOR and the number of linear motifs participating in actual binding events was estimated using the regular expressions listed in the ELM database. In addition, we investigated orthologous proteins present in selected apes and mice.



Our results, based on detailed analysis of the structural/functional features of the proteins, suggest intrinsic disorder is indeed an important feature of postsynaptic proteins and might comprise one of the main factors contributing to the versatility and dynamics of the postsynaptic network. We show that postsynaptic proteins are particularly enriched in disordered segments. Although the number of interacting partner proteins is not exceptionally large, the estimated diversity of the combinations of putative complexes is high in postsynaptic proteins.

Hereinafter, we also look at RNA binding proteins, including those that occur primarily in synapses and postsynaptic density. We analyze their domain diversity and the protein networks they form with each other.

Acknowledgments

The authors acknowledge the support of the National Research, Development and Innovation Office – NKFIH through grants no. NN124363 (to ZG) and the 'János Bolyai Research Scholarship' Program of the Hungarian Academy of Sciences (to ZG).

The research has been carried out within the project Thematic Research Cooperation Establishing Innovative Informatic and Info-communication Solutions, which has been supported by the European Union and co-financed by the European Social Fund under grant number EFOP-3.6.2-16-2017-00013.

References

- [1] W. Feng, M. Zhang, Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nature Reviews Neuroscience* 10, 87–99 (2009).
- [2] M. Zeng, Y. Shang, Y. Araki, T. Guo, R. L. Haganir, M. Zhang, Phase transition in postsynaptic densities underlies formation of synaptic complexes and synaptic plasticity. *Cell* 166, (2016).
- [3] D. Milovanovic, Y. Wu, X. Bian, P. D. Camilli, A liquid phase of synapsin and lipid vesicles. *Science* 361, 604–607 (2018).

PROTEIN NETWORK ANALYSIS

Selected talks

Comparison of the Deterministic and Stochastic Approaches to Michaelis-Menten Kinetics

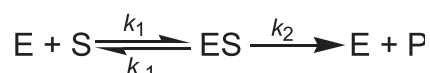
Gábor Lente

Department of General and Physical Chemistry, University of Pécs, Hungary

*E-mail: lenteg@gamma.ttk.pte.hu

Keywords: rate equation, master equation, stochastic maps, single enzyme kinetics

Since it was proposed more than a century ago, the Michaelis-Menten equation [1] has been a standard tool of interpreting kinetic observations in enzyme kinetics [2]. Interestingly, the chemical scheme that leads to this equation and is now commonly called the Michaelis-Menten scheme was in fact postulated by Briggs and Haldane [3]. The scheme involves the reversible formation of an enzyme-substrate adduct (ES) in the second order combination of the free enzyme (E) and the substrate, whereas the product (P) is formed in second, irreversible step:



A textbook derivation employing the quasi-steady state approximation based on this scheme yields the Michaelis-Menten equation:

$$\frac{d[P]}{dt} = \frac{k_2[E]_0[S]}{K_M + [S]}$$

In this equation, K_M is the Michaelis constant, which can be calculated from the rate constants as $K_M = (k_{-1} + k_2)/k_1$. With recent developments, it has become possible to study the activity of a single enzyme molecule [4-5]. As concentrations do not make sense at this single-molecule level, it is clear that the stochastic kinetic approach should be used to interpret the experimental findings in these cases [6]. In this method, a state of a system is identified by giving the molecule numbers. In the studied scheme, the giving number of free enzyme molecules (e) and the number of untransformed substrate molecules (s) is sufficient to identify a state as the other two molecule numbers can be calculated from conservation laws. Then a system of ordinary differential equations is set up for $P_{e,s}(t)$ functions, which give the probability that the system is exactly in state (e,s) at a given time moment t (master equation):

$$\begin{aligned} \frac{dP_{e,s}(t)}{dt} = & - \left[\frac{k_1}{N_A V} es + (k_{-1} + k_2)(e_0 - e) \right] P_{e,s}(t) + \frac{k_1}{N_A V} (e+1)(s+1) P_{e+1,s+1}(t) + \\ & + k_{-1}(e_0 - e + 1) P_{e-1,s-1}(t) + k_2(e_0 - e + 1) P_{e-1,s}(t) \end{aligned}$$

Here e_0 is the overall number of enzyme molecules (free and complexed), which is a conserved quantity, i.e. it does not depend on time. N_A stands for Avogadro's constant, whereas V is the overall volume of the system. Although Arányi and Tóth gave the full symbolic solution for $e_0 = 1$ [7] way before experiments were sensitive enough for measuring the phenomenon itself, this solution seems to have been forgotten and the use of a much more simplistic way of thought has become the norm, the final result of which gives the expectation for the waiting time τ for product formation to occur and is referred as the single-molecule Michaelis-Menten equation [4]:

$$\frac{1}{\langle \tau \rangle} = \frac{k_2[S]}{K_M + [S]}$$

Recent theoretical analysis gave more insight into the molecular-level description of the Michaelis-Menten scheme [6, 8-11]. Particular attention has been given to identifying the region in the phase space of the parameter values (rate constants and initial concentrations) where the use of the stochastic approach is inevitable [8]. These regions can be advantageously presented in graphs that were termed 'stochastic maps'. These maps do not only depend on the parameter values, but also on the target property of interest. To develop these maps, it was also necessary to find the stochastic equivalent of the quasi steady state approach commonly used in deterministic chemical kinetics.

Although the master equation shown above is a linear system of first-order ordinary differential equations with a known general solution, giving this solution in a form that is useful for comparison with experimental results is made practically impossible by the extremely high number of $r_{Pe,s}(t)$ functions necessary to describe the system. To overcome this limitation, a systematic approximation method was developed using a sequence of thought based on combining binomial distributions [9]. It was shown that the time dependent probability of a state can be approximated as:

$$P_{e,s}(t) \approx \binom{w}{e_0 - e} \left((1 - qq(t))^{w - (e_0 - e)} qq(t)^{e_0 - e} \right) \binom{s_0}{s + e_0 - e} p(t)^{s + e_0 - e} (1 - p(t))^{s_0 - s - e_0 + e}$$

The various quantities appearing in this equation are defined as follows:

$$w = \min(e_0, e_0 + s - e)$$

$$qq(t) = \min(q(t)e_0 / w, 1)$$

$$q(t) = \frac{N_A V}{e_0} \lambda_1(t) \frac{1 - \exp(\lambda_2(t)k_1t - \lambda_1(t)k_1t)}{\lambda_1(t) / \lambda_2(t) - \exp(\lambda_2(t)k_1t - \lambda_1(t)k_1t)}$$

$$\lambda_1(t) = \frac{e_0 + p(t)s_0}{2N_A V} + \frac{k_{-1} + k_2}{2k_1} + \sqrt{\left(\frac{e_0 + p(t)s_0}{2N_A V} + \frac{k_{-1} + k_2}{2k_1} \right)^2 - \frac{e_0 p(t)s_0}{(N_A V)^2}}$$

$$\lambda_2(t) = \frac{e_0 + p(t)s_0}{2N_A V} + \frac{k_{-1} + k_2}{2k_1} - \sqrt{\left(\frac{e_0 + p(t)s_0}{2N_A V} + \frac{k_{-1} + k_2}{2k_1} \right)^2 - \frac{e_0 p(t)s_0}{(N_A V)^2}}$$

$$p(t) = \frac{N_A V (k_{-1} + k_2)}{s_0 k_1} W \left(\frac{k_1 s_0}{N_A V (k_{-1} + k_2)} \exp \left(\frac{k_1 (s_0 - e_0 k_2 t)}{N_A V (k_{-1} + k_2)} \right) \right)$$

In the last formula, W stands for the Lambert W function, which is the inverse of the $x \exp(x)$ function and appears in the solution of the deterministic Michaelis-Menten equation as well.

Acknowledgments

The research was financed by the Higher Education Institutional Excellence Program of the Ministry of Human Capacities in Hungary, within the framework of the 1st thematic program of the University of Pécs.

References

- [1] L. Michaelis, M. L. Menten. Die Kinetik der Invertinwirkung. Biochemische Zeitung, 1913, 49, 333-369
- [2] A. Cornish-Bowden. One hundred years of Michaelis-Menten kinetics. Perspectives in Science, 2015, 4, 3-9
- [3] G. E. Briggs, J. B. S. Haldane. A note on the kinematics of enzyme action. Biochemical Journal, 1925, 19, 338-339
- [4] S. C. Kou, B. J. Cherayil, W. Min, B. P. English, X. S. Xie. Single-molecule Michaelis-Menten equations. Journal of Physical Chemistry B, 2006, 109, 19068-19081

- [5] B. P. English, W. Min, A. M. van Oijen, K. T. Lee, G. Luo, H. Sun, B. Cherayil, S. C. Kou, X. S. Xie. Ever-fluctuating single enzyme molecules: Michaelis–Menten equation revisited. *Nature Chemical Biology*, 2006, 2, 87-94
- [6] P. Érdi, G. Lente. *Stochastic Chemical Kinetics. Theory and (Mostly) Systems Biological Applications*. Springer, 2014
- [7] P. Arányi, J. Tóth. A full stochastic description of the Michaelis–Menten reaction for small systems. *Acta Biochimica et Biophysica Academiae Scientificarum Hungariae*, 1977, 12, 375-388
- [8] É. Dóka, G. Lente. Stochastic mapping of the Michaelis-Menten mechanism. *Journal of Chemical Physics*, 2012, 136, 054111
- [9] G. Lente. A binomial stochastic kinetic approach to the Michaelis-Menten mechanism. *Chemical Physics Letters*, 2013, 568-569, 167-169
- [10] S. Bankó, Z. Kucsma, G. Lente, C. Bagyinka. The autocatalytic step is an integral part of the hydrogenase enzyme-cycle. *Biochimica and Biophysica Acta - Proteins and Proteomics*, 2013, 1834, 658-664
- [11] G. Lente. A Michaelis-Menten-kinetika determinisztikus és sztochasztikus modelljei. *Alkalmazott Matematikai Lapok*, 2016, 33, 159-174 (Hungarian)

PROTEIN NETWORK ANALYSIS

Selected talks

Ensemble-based interpretation of the dynamic allostery in a PDZ domain

Anett Hinsenkamp^{1*}, Dr. Zoltán Gáspári²

¹3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

²Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, Budapest, 1083, Hungary

*E-mail: hinsenkamp.anett@itk.ppke.hu

Keywords: PSD-95, dynamic simulation, structural ensemble

The postsynaptic density (PSD) of excitatory synapses is a highly organized, dynamic network of proteins responsible for the regulation, trafficking and anchoring of the membrane receptors. Therefore, these scaffold proteins are involved in the processes of learning, memory and synaptic plasticity.[1,2] Atomic-level understanding of the structure-dynamics relationships, including allostery in partner binding in these proteins is crucial in building detailed models of synaptic transmission. PDZ domains, specialized for protein-protein interactions, serve as nodes in multiple PSD scaffold protein connections. PDZ domains (named after the first three proteins discovered to share the domain: PSD-95, Discs-large, ZO-1) are around 90 residue long globular protein modules consisting of five to six β -sheets and 2 α -helices with a binding pocket between the second β -sheet and the second α -helix. PDZ domains typically interact with the C-terminus of various proteins and peptides, but other forms of interactions are also known.[1-3] They are commonly classified according to the C-terminal sequences of their binding partners, although the specificity determinants of the interaction partners are more complex. Structured extensions or additional elements (helices, loops, β -strands) can be found in a number of PDZ domains, which may influence binding.[4] For example, the third PDZ domain (PDZ3) of PSD-95 (postsynaptic density protein 95) has an additional helix at its C-terminus.[5] Andrew Lee and coworkers showed that the additional helix of PSD-95 PDZ3 affects the binding of CRIPT peptide even though it is not in direct contact with the binding site. Truncation of the third helix leads to an over twenty fold decrease in the binding affinity, where the difference in free energy is entropic in origin.[6] Phosphorylation of conserved tyrosine 397 in the third helix occurs in vivo implying a possible biological regulatory mechanism of the additional helix. [7]

Detailed, atomic-level structural models were generated to reflect the experimentally determined mobility of the free and ligand-bound states of PDZ3 variants. These models can help in establishing the mechanistic details behind the allosteric behavior of PDZ3 and characterize the conformational changes. [8,9] Multiple structural ensembles were generated using molecular dynamics simulations with NMR derived mobility parameters as restraints.[6] Free and complexed forms of PDZ3 and CRIPT were constructed and the C-terminal truncation variants ($\Delta 7$ ct) of these. Since the transition between these states take place on the millisecond time scale, only models of the inherent dynamics were created and compared. The use of NMR-derived parameters as restraints enables both the structure and nanosecond time scale dynamics to be represented. Lipari-Szabo model-free S2 order parameters obtained from NMR spin relaxation experiments were added to the simulations as restraints. The S2 order parameter describes how restricted the motion of the given substructure is.[10] The comparison of ensembles restrained with only backbone, only side-chain and both types of S2 order parameters revealed the importance of the use of side-chain mobility data next to backbone mobility parameters. Ensemble restrained simulations of different lengths were performed in order to map the conformational space visited along motions on the fast timescale. Entropy estimates of the ensembles concur with experimental observations described in the literature.[5] The high entropy difference upon binding in the PDZ variant lacking the C-terminal extension

is in correspondence with the increased side-chain mobility of the corresponding free form. From principal component analysis of the ensembles, backbone conformational differences can be observed between the free and full states. From the side-chain and backbone mobility results and possible hydrogen bonding patterns, key regions of the domain, involved in binding and the connection between the binding site and the additional helix were identified. Our results point to the existence of a set of residues whose properties are in tight connection with the alterations of internal dynamics of the PDZ3 domain upon ligand binding.

Our investigations reveal the importance of atomic-level structural models that show a reasonable quantitative agreement with available experimental data. Our preliminary results confirm the importance of including both side-chain and backbone dynamic parameters in representations. Our models allow the identification of key residue contacts and their rearrangement upon ligand binding and interactions with the additional C-terminal helix and allow the design of targeted experiments to investigate the role of individual residues and residue-residue interactions. The modeled structures also provide a first step towards high-resolution modeling of the PSD network and its possible rearrangements during molecular processes of learning and memory.

Acknowledgments

The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] E. Kim, M. Sheng. PDZ domain proteins of synapses. *Nature Reviews Neuroscience*, 2004, 5, 771-781
- [2] W. Feng, M. Zhang. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nature Reviews Neuroscience*, 2009, 10, 87-99
- [3] S. Steiner, A. Cafisch. Peptide binding to the PDZ3 domain by conformational selection. *Proteins*, 2012, 80, 2562-2572
- [4] S. Mostarda, D. Gfeller, F. Rao. Beyond the binding site: The role of the β 2- β 3 loop and extra-domain structures in PDZ domains. *PLOS*, 8, 3
- [5] D. A. Doyle, A. Lee, J. Lewis, E. Kim, M. Sheng, R. MacKinnon. Crystal Structures of a Complexed and Peptide-Free Membrane Protein-Binding Domain: Molecular Basis of Peptide Recognition by PDZ. *Cell*, 1996, 85, 1067-1076
- [6] C. M. Petit, J. Zhang, P. J. Sapienza, E. J. Fuentes, A. L. Lee. Hidden dynamic allostery in a PDZ domain. *PNAS*, 2009, 106, 18249-18254
- [7] J. Zhang, C. M. Petit, D. S. King, A. L. Lee. Phosphorylation of a PDZ Domain Extension Modulates Binding Affinity and Interdomain Interactions in Postsynaptic Density-95 (PSD-95) Protein, a Membrane associated Guanylate Kinase (MAGUK). *Journal of Biological Chemistry*, 2011, 283, 41776-41785
- [8] A. F. Ángyán, Z. Gáspári. Ensemble-Based Interpretations of NMR Structural Data to Describe Protein Internal Dynamics. *Molecules*, 2013, 18, 10548-10567
- [9] A. Dhulesia, J. Gsponer, M. Vendruscolo. Mapping of two networks of residues that exhibit structural and dynamical changes upon binding in a PDZ domain protein. *Journal of the American Chemical Society*, 2008, 130, 8931-8939
- [10] G. Lipari, G. Szabó. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. *Journal of the American Chemical Society*, 1982, 104, 4546-4559

PROTEIN NETWORK ANALYSIS

Selected talks

Reachability analysis of discrete state chemical reaction networks obeying conservation laws

Gergely Szlobodnyik^{1,2,*}, Gábor Szederkényi¹

¹Pázmány Péter Catholic University Faculty of Information Technology and Bionics

²in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: szlger91@gmail.com

Keywords: reachability analysis, discrete state space, chemical reaction networks, conservation laws, Markov chains

There are several different modeling approaches that are widely used in characterizing the dynamical behavior of complex networks of chemically interacting components (species). Deterministic ordinary differential equation systems based models are one of the most common modeling choice in mathematical and systems biology. The continuous state modeling approach poses strong assumptions to the underlying system. It assumes high molecular count of the interacting species and their homogeneous (well-stirred) distribution in the media. However, in several systems of bio(chemical) importance (e.g. genetic networks, intracellular signal transduction pathways), the molecular count of the components are known to be low (e.g. < 100 molecules per species), hence the assumptions for deterministic modeling do not hold and it becomes necessary to introduce a discrete state mathematical modeling approach to keep track the molecular counts of the different species of the system. To model the dynamical behavior of discrete state systems Markov chains are commonly employed [1].

One of the primary interests of Chemical Reaction Network Theory (CRNT) is to examine the relations between the qualitative dynamical behavior and the underlying network structure of complex (bio)chemical reaction networks consisting of several interacting species [2]. To do so it is important to make use of properties of the chemical reaction network (CRN) that are related to the qualitative dynamical behavior and the network structure, as well. In the context of discrete state chemical reaction networks (d-CRNs) the reachability analysis can be applied to study the qualitative dynamical behavior and the network topology at the same time. Reachability analysis poses the question of whether is it possible to reach a prescribed final state from a prescribed initial one, given the underlying network topology? It is known that the reachability relation between any pair of non-negative initial and target states is uniquely determined by the underlying network topology itself, but it has significant effect on the qualitative dynamical behavior as well.

By means of reachability analysis several problems can be efficiently expressed in a formal mathematical way and analyzed. One of them having particularly high interest in the CRNT literature is the existence of extinction events, i.e. the existence of trajectories leading to the (irreversible) extinction of species from the system. It is known that subconservative d-CRNs under some conditions on the network structure are guaranteed to exhibit an extinction event from any point of the state space [3, 4].

The formal mathematical model of d-CRNs is known to be equivalent to a well-known discrete state model of theoretical computer science, the so-called vector addition systems with states (VASS) or equivalently Petri nets. This implies that the d-CRN reachability problem is equivalent to the extensively studied problem of Petri net reachability. The Petri net reachability problem is known to be decidable, but beyond some special structural subclasses, in general case it is not known whether there exists an algorithm of primitive-recursive time complexity deciding the reachability problem on an arbitrary network topology and prescribed initial and

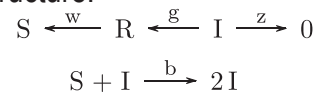
final states [5, 6, 7, 8].

In this talk we consider the reachability problem of chemical reaction networks obeying conservation laws, namely sub- and superconservative CRNs. The reachable state space of subconservative CRNs is known to be bounded, while that of a superconservative CRN is unbounded [9]. We trace back the reachability problem of superconservative CRNs to the subconservative one for which the reachable state space is guaranteed to be bounded. Hence in the sequel it is enough to restrict ourselves to the class of subconservative systems with respect to reachability problems.

Then the reachable state space of subconservative d-CRNs is geometrically characterized by well-defined simplexes. Upper bounds are also given to the maximal length of cycle-free admissible state transitions sequences (state space trajectories) associated to a given reachability problem [10].

The general subconservative d-CRN reachability problem is reformulated as an Integer Programming (IP) feasibility (decision) problem. Making use of the Lenstra algorithm it is shown that for any subconservative CRNs the reachability problem has polynomial time complexity, given the fixed dimension of the problem [11]. We also identified d-CRN subclasses for which the reachability is equivalent to the existence of a non-negative integer solution of the d-CRN state equation. It is shown by a constructive proof that for any d-CRN of stoichiometric matrix with rank smaller than or equal to 2, the non-negative integer solution of the state equation implies the reachability. Then this result is extended to the case of sub- and superconservative d-CRNs of arbitrary high dimensional state space composed of reactions having at most one source and at most one product species beyond possible catalyzer species. This way the reachability problem is reformulated in the form of an IP feasibility problem which is solvable in polynomial time in the number of constraints and in the distance of the initial and target states [10].

As an illustrative example we consider the SIRS (susceptible-infected-recovered-susceptible) epidemiological model of the following reaction network structure:



Note that 0 denotes the zero-complex and the respective transition $I \rightarrow 0$ represents mortality. The state variable $X \in \mathbb{Z}_{\geq 0}^3$ tracks the number of individuals in the compartments in a fixed order: susceptible (S), infected (I) and recovered (R). Using the proposed IP feasibility based computational approach it can be shown that the discrete state SIRS model can result in an extinction event. We set the initial and target states of the system to be $X_0 = [50 \ 2 \ 0]^T$ and $X' = [32 \ 0 \ 0]^T$, respectively, where the target state X' represents an extinction because the number of infected individuals is equal to zero. The reachability of X' from X_0 is validated. By equipping the feasibility problem with the cost function $\sum_{i=1}^l [c]_i$ we determined a minimal-length state transition sequence as depicted in Figure 1.

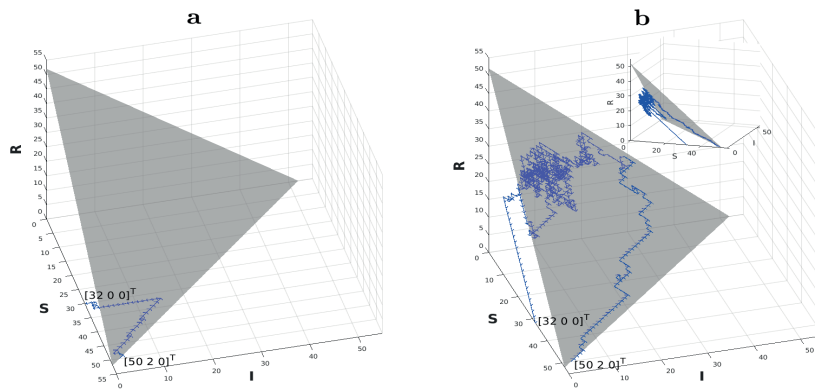


Figure 1: State transition sequences reaching the same target state $[32 \ 0 \ 0]^T$ where the disease vanishes.

a) the shortest path determined by IP, **b)** a sample path simulated by the next reaction method.

Acknowledgement

The research has been supported by the European Union, co-financed by the European Social Fund through the grant EFOP-3.6.3-VEKOP-16-2017-00002.

The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] D. F. Anderson, T. G. Kurtz, "Stochastic analysis of biochemical systems", Berlin Springer, 2015.
- [2] D. Angeli, "A tutorial on chemical reaction network dynamics", *European Journal of Control*, 15., 2009., 398-406.
- [3] D. F. Anderson et. al., "Stochastic analysis of chemical reaction networks with absolute concentration robustness", *J. R. Soc. Interface*, 11(90), 2014.
- [4] M. D. Johnston, et. al., "Conditions for extinction events in chemical reaction networks with discrete state spaces", *J. Math. Biol.*, 2017. [5] E. W. Mayr, "An algorithm for the general petri net reachability problem", *STOC 1981*, pp. 238–246. ACM, 1981
- [6] S. R. Kosaraju, "Decidability of reachability in vector addition systems", *STOC 1982. ACM*, pp. 267–28 [7] J. L. Lambert, "A structure to decide reachability in Petri nets", *Theor. Comput. Sci.* 99(1), pp. 79–104, 1992
- [8] S. Schmitz, "Complexity Hierarchies beyond Elementary", *ACM Transactions on Computation Theory*, 8(1), pp. 1-36. , 2016.
- [9] G. Memmi, G. Roucairol, "Linear algebra in net theory", In W. Brauer (ed.), "Net Theory and Applications", vol. 84. of *Lecture Notes in Comput. Sci.*, pp. 213–223. Springer, 1975.
- [10] G. Szlobodnyik, G. Szederkenyi, M.D. Johnston, "Reachability Analysis of Subconservative Discrete Chemical Reaction Networks", accepted in *MATCH Commun. Math. Comput. Chem.*
- [11] H. W. Lenstra, "Integer Programming with a Fixed Number of Variables", *Math. Oper. Res.*, 8, pp. 538–548., 1983

The background is a light blue grid with various geometric shapes. At the top, there are faint, light blue hexagonal outlines with small dots at their vertices. Below these, there are several solid blue hexagons of different sizes and orientations. Some of these hexagons are interconnected by thin white lines, forming a network-like structure. The overall aesthetic is clean, modern, and technical.

POSTER PRESENTATIONS

for the two Conference

POSTER SESSION

Effect of protein aggregation on protein complex assemblies

Rita Szilvia Kiss^{1,2*}, Bence Keömley-Horváth^{1,2*}, Simone Rizzetto^{3,4*}, Attila Csikász-Nagy^{1,2,5*}

¹Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary

²3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

³The Microsoft Research - University of Trento Centre for Computational and Systems Biology, Rovereto, Italy

⁴School of Medical Sciences, UNSW Australia, Sydney, Australia

⁵Randall Centre for Cell and Molecular Biophysics, King's College London, London, UK

*E-mail: kissrita.szilvia@gmail.com (RSK); horvathb94@gmail.com (BKH); simone.rizzetto@outlook.com (SR);

attila.csikasz-nagy@itk.ppke.hu (ACN)

Keywords: protein complex, protein aggregation, neurodegeneration

Although protein complexes and structures are widely researched, most of them are still undiscovered. To comprehend cellular regulatory mechanisms and support new drug development, we have to determine these complexes. Accordingly, getting detailed information about the abundance and composition of protein complexes is essential. A new approach is being proposed with a software called SiComPre (Simulation based Complex Prediction), which makes stochastic simulations of protein complex formation. It integrates multi-source data, for example protein abundances and domain-domain interactions, combines these information, and generates a computational model of all proteins and their dynamic interactions. Different cell-specific datasets are available to predict protein complexes specific to budding yeast or human. Using these datasets various stochastic simulations are performed in order to predict protein complexes specific to cells in different species. Also, the results provide an approximation of protein complex abundances.[1] Simulations in the software is also able to predict how composition and abundance of protein complexes can be influenced by perturbations of a drug. Since SiComPre predicts alternative protein complex forms and their abundances, and can also be used for complexome change predictions during drug treatment, it could be used for qualitative and quantitative difference predictions in the complexome in numerous tissue types and conditions if additional data is provided.[1][2]

The method can be used to alter protein complexes when certain proteins are aggregating. Such protein aggregation can cause neurodegenerative diseases and the expression of proteins that are important in neurodegenerative diseases, such as alpha-synuclein, is toxic, also in yeast cells.[3] For these yeast experiments proteomic data series are available for several proteins, for instance alpha-synuclein and amyloid-beta, including the ratio of protein abundance in perturbed and control cells. [4]

To simulate protein complexes properly, the most accurate data has to be provided as an input. The current input is not accurate enough, so by using the proteomic data provided by the yeast experiments it is possible to improve and optimize our input. By using SiComPre's simulation the attempt to find out which changes in protein complexes can lead to cell death in neurodegenerative diseases is probable.

Acknowledgement

The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and infocommunications)

References

- [1] Rizzetto, Simone, Corrado Priami, and Attila Csikász-Nagy. "Qualitative and quantitative protein complex prediction through proteome-wide simulations." *PLoS computational biology* 11.10 (2015): e1004424.
- [2] Rizzetto, Simone, et al. "Context-dependent prediction of protein complexes by SiComPre." *NPJ systems biology and applications* 4.1 (2018): 37.
- [3] Ross, C. A., & Poirier, M. A. (2004). Protein aggregation and neurodegenerative disease. *Nature Medicine*, 10(7), S10–S17.
- [4] Andre Melnik, Dissertation (2015), Comparative analysis of intracellular responses induced by amyloidogenic proteins

Dynamical modeling of the most important pathways in cellular metabolism

Gergely Sváb^{1*}, László Tretter¹, Gábor Szederkényi²

¹Department of Medical Biochemistry, Semmelweis University, Budapest, Hungary

²Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary

³in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: svab.gergely@med.semmelweis-univ.hu

Keywords: dynamical modeling, metabolism, molecular networks, computer simulation

Recently, the study of cell metabolism has been in the focus of biological research, since most of the changes in cell function cause metabolic changes. Observation of these changes is usually difficult, but there are promising possibilities in metabolic pathway modeling. There is a clear need for cellular models which can predict the metabolic states and modified enzyme activity in modified states. Mitochondria which are cell organelles with prokaryotic origins have a key role in cell metabolism. Degradation of carbohydrates, fats and proteins takes place in the mitochondria, and the intermediers formed may partake in anabolic processes.

In our previous works we examined the intermediers of the citric acid cycle. These metabolites can transform to each other according to the actual metabolic state in mitochondria. The key elements in the regulation of citric acid cycle are the concentrations of intermediers and the changes in redox potential. The initial model was extended by adding mitochondrial transporters which connect cellular and mitochondrial metabolic networks. The model can describe the mechanism of competitive inhibition and the function of multifunctional enzymes which play an important role in the metabolism of several different substrates.

In this work we present the model complemented with several mitochondrial transporters and cellular degradation pathways. For example, glycolysis pathway contains the metabolism of glucose and fructose and leads to pyruvate. Several amino acid metabolic pathways are included, which leads to ammonia, which is then used in the urea cycle. Using the constructed model, we analyze the temporal changes of the metabolite concentrations with a quantitative mathematical description.

Acknowledgments

The research was partially supported by the grant K115694 of the National Research, Development and Innovation Office - NKFIH. The project has also been supported by the European Union, co-financed by the European Social Fund through the grants EFOP-3.6.2-16-2017-00013 and EFOP-3.6.3-VEKOP-16-2017-00009.

References

- [1] Wu F, Yang F, Vinnakota KC, Beard DA. Computer modeling of mitochondrial tricarboxylic acid cycle, oxidative phosphorylation, metabolite transport, and electrophysiology. *The Journal of biological chemistry*, 282:24525–24537, August 2007.
- [2] Nazaret C, Heiske M, Thurley K, Mazat JP. Mitochondrial energetic metabolism: a simplified model of tca cycle with atp production. *Journal of theoretical biology*, 258:455–464, June
- [3] Korla K, Mitra CK. Modelling the krebs cycle and oxidative phosphorylation. *Journal of biomolecular structure & dynamics*, 32:242–256, 2014.
- [4] Korla K, Vadlakonda L, Mitra CK. Kinetic simulation of malate-aspartate and citrate-pyruvate shuttles in association with krebs cycle. *Journal of biomolecular structure & dynamics*, 33:2390–2403, 2015.
- [5] Korla K. Reactive oxygen species and energy machinery: an integrated dynamic model. *Journal of biomolecular structure & dynamics*, 34:1625–1640, August 2016.

Towards a systematic census of bacterial quorum sensing genes in public databases

János Juhász^{1,2*}, Éva M. Kistóth^{1,2}, Z. Nagy¹, P. Szolgay¹, Sándor Pongor¹ and Balázs Ligeti^{1,3}

¹Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary;

²3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

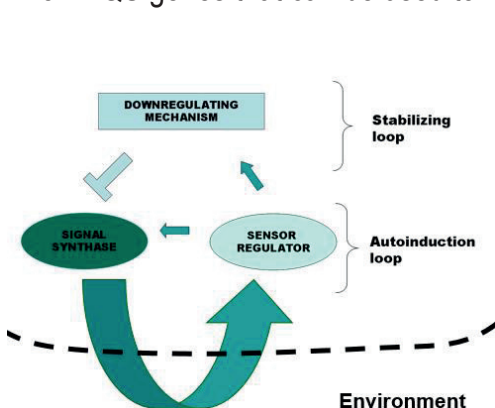
³Institute of Medical Microbiology, Semmelweis University, Budapest, Hungary MTA TTK

*E-mail: juhasz.janos@itk.ppke.hu

Keywords: bacterial signaling, quorum sensing, operons

Summary

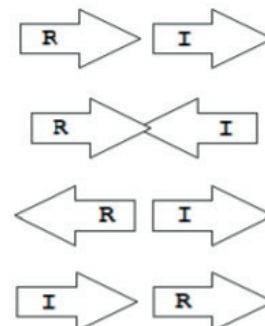
The growing body of sequence data offers new possibilities to find new bacterial quorum sensing (QS) genes. Here we outline a method that allows one to extract sequence as well as chromosomal position patterns from known QS genes that can be used to find similar gene arrangements in unannotated sequence data.



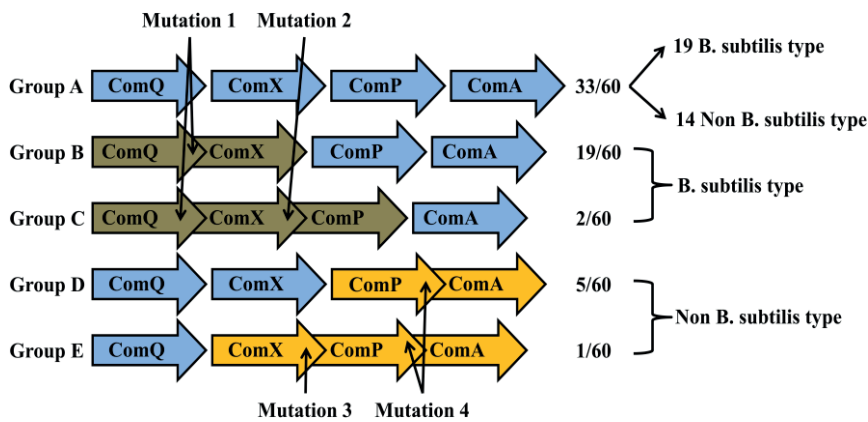
Quorum sensing signaling is an autocrine signaling mechanism present in various unicellular organisms, mainly in bacteria [1]. This mechanism is based on an incoherent feed forward network that, by definition consists of a positive feedback and negative feedback control loop. The heart of this mechanism is the positive feedback or autoinducer loop based on an enzymatic production of a signal molecule that binds to a sensor-receptor molecule that upregulates the production of the signal as well as the expression of various other genes. The “trick” of this mechanism is that the signal molecule can leave from as well as return to the cell either by active or by passive transport. The signal molecules outside the cell can influence the

metabolism of other cells, so there will be a de facto communication between cells. If signal concentration in the extracellular space will be high enough, gene expression in all concerned cells will be upregulated, so the functioning of the cell population will be synchronized. This simple phenomenon allows a cell population to solve problems that individual cells cannot tackle, such as the colonization of a surface or infecting a host organism. On the other hand, the negative feedback loop plays a stabilizing role that does not allow signal production to grow without limits.

As an example, luxR genes encode a transcriptional regulator that control acyl homoserine lactone-based quorum sensing (AHL QS) in many Gram negative bacteria. In this system the AHL signal travels through the bacterial cell wall via passive transport. On the bacterial chromosome, luxR genes are usually in the direct vicinity of a luxI gene encoding the AHL signal synthase. Genes involved in stabilization are often in between these two genes or are located next to them. Over 15 operon types were observed in the AHL signaling family alone [2-3]. Another well studied QS mechanism in Gram positive bacteria is the comQXPA locus in *Bacillus subtilis* and related bacteria which encodes a QS system consisting of 4 genes. Here the signal is a peptide that is transported across the membrane via active transport. The peptide in the extracellular space is sensed by the extracellular part of a transmembrane receptor ComP. The intracellular part of the receptor is a histidine kinase which will phosphorylate a DNA-binding protein ComA. Once phosphorylated, ComA will bind to the



chromosome and upregulate the production of a ComX protein that includes the peptide signal. This protein will then be cleaved by a transmembrane protein ComQ that will pump out the peptide signal into the extracellular space. The autoinducer loop of this system thus consists of 4 proteins and includes active transport, in contrast to the AHL system where the autoinducer loop consists of only two proteins and is based on passive transport. On the other hand, the topologies of the comQXPA genes are quite conserved, minor differences occur only in the overlap of the concerned genes [4].



A preliminary overview of the current literature revealed about 20 further well studied quorum sensing systems. Comprehensive sequence collections were published on the AHL and the comQXPA systems approximately 5 years ago, but the body of available bacterial sequences has grown about 10 fold in the meantime so a survey of new data is an important task. The challenge of such a survey is the variability of the

QS systems. Importantly, we can safely detect only the known, and let's add, well known QS systems. The strategy tries to generalize the logics of our previous surveys, i.e. a QS system is considered as a generalized structure of entities and relationships which constitutes a graph in which protein coding genes are the nodes and intergenic distances are the edges. In order to detect such a gene set within the chromosome we need to apply parsimonious and scalable solutions because the number sequences to be screened is now many millions and the number is exponentially growing.

One of the problem complicating the situation are the truncated topologies, i.e. operons where one or more members are missing. For instance, AHL operons contain only two genes, luxI and luxR but many bacteria contain solo luxR genes, i.e. receptor genes with no signal synthase. Solo luxR genes are even more frequent than complete AHL operons. This can have many reasons. For instance, the survey did not pick up the adjacent luxI homolog either because its sequence is too divergent or because it is further away within the chromosome. Or the luxR gene may control an unknown type of signal synthase (as it was found in a few cases). Or the solo luxR protein responds to an unknown type of a signal. With more complex operons the situation is even more complicated, so we designed a search algorithm that employs a hierarchy of search space reduction steps which is based on a hierarchy of molecular descriptions, namely presence-absence, composition and full structure descriptions. For instance, if we have an operon of four members, we will keep only genomes where at least three of the components are present. From these we concentrate on genomes where the required number of elements are present within a certain distance, a value observed in the known instances of the operon. Finally establish the gene distances and write down the topology found. This is a highly efficient space reduction strategy since the compute/intensive steps are limited to very few cases.

An additional challenge is the recent surge in the number of next generation sequencing data obtained on various bacterial systems. Such data, such as those present in the NCBI SRA archive consist of many million reads each, and it is important to know whether or not QS systems are present or are active in them. The collection of QS genes developed within our project will be a useful tool for detecting such genes directly from reads and from metagenomics datasets.

The long term goal of this project is develop automated protocols to extract QS genes from genomic data. The output would be the topological description of the QS operons along with a quality indicator characterizing the reliability of the prediction. For the predictions we will use Hidden Markov Models as well as fast sequence compression programs (Bowtie or BWA) ported to multicore architectures such as GPU and FPGA which will be used as search engines of dedicated web servers.

Acknowledgments

Funded in part by Hungarian government grants OTKA 120650, EFOP-3.6.2-16-2017-00013 (supported by the European Union and co-financed by the European Social Fund) and ED_17-1-2017-0009 (within the National Bionics Program)

References

- [1] M. Whiteley , S.P. Diggle, E.P. Greenberg: Progress in and promise of bacterial quorum sensing research. *Nature*, 2017 Progress in and promise of bacterial quorum sensing research. *Nature*. 2017, 551, 313-320.
- [2] Z. Gelencsér, K.S. Choudhary, B.G. Coutinho, S. Hudaiberdiev, B. Galbáts, V. Venturi, and S. Pongor, S. Classifying the topology of AHL-driven quorum sensing circuits in proteobacterial genomes. *Sensors*. 2012, 12, 5432-5444.
- [3] K.S. Choudhary, S. Hudaiberdiev, Z. Gelencsér, B.G. Coutinho, V. Venturi, S. Pongor, The Organization of the Quorum Sensing luxI/R Family Genes in Burkholderia. *Int. J. Mol. Sci.* 2013, 14 13727-13747.
- [4] I. Dogsa, K.S. Choudhary, Z. Marsetic, S. Hudaiberdiev, R.Vera, S. Pongor, S. and I. Mandic-Mulec, I. ComQXPA quorum sensing systems may not be unique to *Bacillus subtilis*: a census in prokaryotic genomes. *PlosOne*, 2014, 9 e96122.

Network-based prediction of drug-combination efficiency based on expression data on off target genes

Balázs Ligeti^{1,2*}, Otilia Menyhárt^{3*}, János Juhász^{1,6}, Sándor Pongor¹ and Balázs Györfy^{3,4,5}

¹Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary;

²Institute of Medical Microbiology, Semmelweis University, Budapest Hungary MTA TTK

³Lendület Cancer Biomarker Research Group, Budapest, Hungary

⁴2nd Department of Pediatrics, Semmelweis University, Budapest, Hungary,

⁵MTA-SE Pediatrics and Nephrology Research Group, Budapest, Hungary

⁶3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

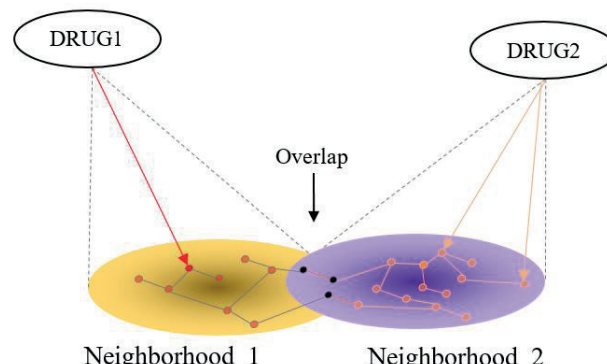
*Equal contributions. E-mail: Balazs Ligeti <ligeti.balazs@itk.ppke.hu>, Otilia Menyhárt <vicabuba@gmail.com>

Keywords: drug-combinations, off target, bioinformatics, prediction, complex diseases

Summary: Gene expression data from public databases can be efficiently incorporated into algorithms predicting drug combination efficiency by using a generalized framework of network neighborhoods. In this representation, efficient drugs are those where the affected targets overlap well with the known disease genes, and synergistic drugs are those whose neighborhoods overlap well with the purported disease genes.

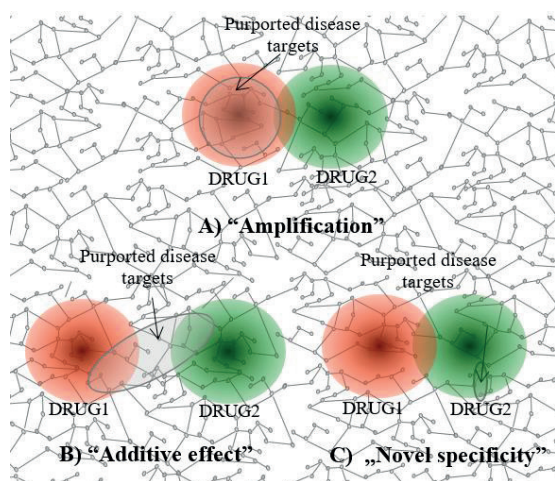
Drug combinations are very efficient and frequently used tools in the treatment of complex multigene diseases such as hypertension, diabetes and cancer [1]. Broadly speaking, the effect of combining pharmacons can be either synergistic (beneficial), or additive (no interactions) or antagonistic (negative, adverse) and delineating the effect of combination therapies must include careful scrutiny of these possibilities [2]. Nevertheless, most currently used combinations were found in tedious empirical ways, consequently there is

a growing interest for finding potentially interacting drug partners using in silico methods. One of the potential avenues is the network propagation hypotheses [3,4] which posits that the genes or proteins, included those affected by pharmacons, are interlinked by an interaction network, which is perturbed by the action of drugs.



It is supposed that perturbations can propagate (diffuse) via the links of this network and combination effects will result in unexpected accumulation of perturbations at various points of the network. The gross effect of drug combination as the portion of jointly perturbed genes which gives a simple [0,1] measure, the so called Target Overlap Score (TOS), TOS was shown to correlate well both with known drug/combination data and with the results of newly conducted clinical trials. However there are also limitations, such as the lack of distinction between positive and negative correlations, and the difficulty of introducing experimental data into the system.

In this work we plan to integrate experimental data available in public databases into network based prediction of drug interactions. The idea is to use a generalized framework of



network interactions in which both pharmacon target genes and disease-affected genes are symbolically represented as off-target gene neighborhoods. Known off-target genes can also be added as separate neighborhoods. In the previous studies, the neighborhoods were computationally generated by calculating the diffusion of perturbation around a known target or off-target gene. Here we make use of experimentally determined neighborhoods extracted from a database of gene expression data, mainly microarray experiments. It is also possible to combine the two principles (ie. using calculated and experimentally derived neighborhoods) by calculating the diffusion neighborhood around experimentally validated target genes. This generalization may allow one to incorporate

Validation of the new principle can be carried out by using public data taken from the DCDB, TTD and Drugs.com databases, or using published clinical data. Preliminary studies show that incorporating gene expression data improves the correlation, so we trust that this principle can further improve the efficiency of predictive algorithms.

Acknowledgments

Funded in part by Hungarian government grants OTKA 120650, EFOP-3.6.2-16-2017-00013 (supported by the European Union and co-financed by the European Social Fund) and ED_17-1-2017-0009 (within the National Bionics Program)

References

- [1] Analysis of drug combinations: current methodological landscape. *Pharmacol Res Perspect.* 2015, e00149
- [2] F. Schmidt F, H. Matter, H. Hessler and A. Czich Predictive in silico off-target profiling in drug discovery. *Future Med Chem.* 2014, 6, 295-317
- [3] B. Ligeti, Zs. Pénczvártó, R. Vera B. Györfy, S. Pongor A Network-Based Target Overlap Score for Characterizing Drug Combinations: High Correlation with Cancer Clinical Trial Results, *PlosOne*, 2015, 10, e0129267
- [4] B. Ligeti, O. Menyhárt, I. Petrič B. Györfy, S. Pongor Menyhárt O, Petrič I, Györfy B, Pongor S. Propagation on Molecular Interaction Networks: Prediction of Effective Drug Combinations and Biomarkers in Cancer Treatment. *Current Pharmaceutical Design.* 2017, 23, 5-28,.

POSTER SESSION

Modelling genomic and metabolic responses of microbial communities to antibiotics

János Juhász^{1,2*}, Balázs Ligeti^{1,3}, and Sándor Pongor¹

¹Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary;

²In Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

³Institute of Medical Microbiology, Semmelweis University, Budapest Hungary

*E-mail: juhasz.janos@itk.ppke.hu

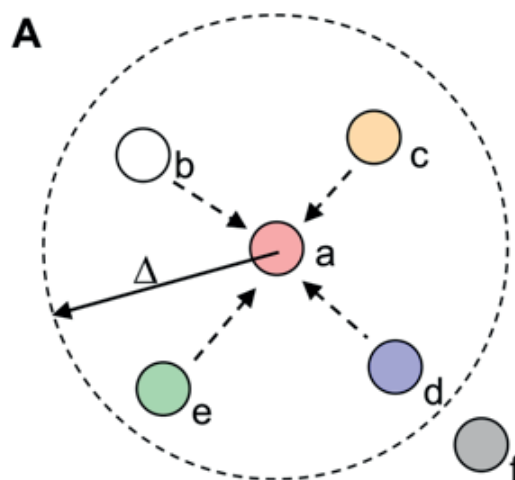
Keywords: microbiome, quorum sensing, horizontal gene transfer

Summary

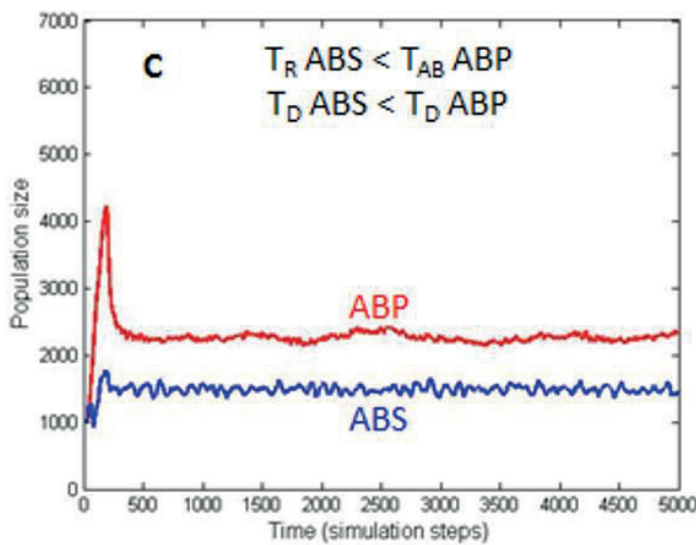
Multispecies bacterial communities can respond to antibiotic challenge both by genomic and by metabolic changes. Agent based modelling shows that in both cases, multiple antibiotic factors will be present in the community which will provide a kind of territorial defense, i.e. external bacteria will have difficulties invading the community.

Multispecies bacterial communities, though not always visible to human eyes, represent the largest metabolic potential present in nature. They can be found as large microbial mats on the sea-floor, as gut microflora of warm blooded animals and humans, etc. They are also major contributors to nature's evolution since they were the predominant life form before foraging animals appeared, and traces of these periods can be seen today as limestone layers of current mountain ranges are believed to be the calcified forms of such microbial mats. Today, of special, medical interest are the microbiota present in the human body and in the human environment, as the emergence of resistance due to the continuing massive application of antibiotic drugs is one of the key public health challenges.

The current project was undertaken in order to get insights into the effects of antibiotics on the stability of microbial communities using computer modelling. Since this is a purely theoretical approach we use a generalized definition for antibiotic action, i.e. any molecule that negatively affects the growth of another bacterium is considered an "antibiotic". In other words, this definition also includes antimicrobials and bacteriostatic agents, not only *lege artis* antibiotics manufactured by the pharmaceutical industry. Second, we are using agent based models as the major modelling tools. In this representation, bacteria are the agents that carry out computations at given time steps. Third we are interested in two kinds of community responses. Evolutionary (or genetic) responses involve changes to the genome via horizontal gene transfer; bacteria acquire antibiotic production and resistance genes from the environment, which requires several generations. In contrast, metabolic responses are immediate, bacteria can turn on the production of, say, enzymes capable of digesting the antibiotics present in the environment thereby helping the survival of both themselves as well as other species present in the community. We showed that horizontal transfer of resistance genes may contribute to the formation of stable and diverse communities capable of protecting themselves with a battery of antimicrobial agents while preserving a varied



metabolic repertoire. In other words horizontal transfer of resistance genes makes a community compatible in terms of exoproducts and fence off the attack of intruders that cannot adapt to all antimicrobials present within the community [1].



Metabolic adaptation was modelled in the context of microbial communities that share the same secreted signals that turn on certain metabolic activities. Here we showed that differential signal-response thresholds provide a simple way for keeping slower-growing species within the community. High response thresholds can be considered as self-restraint of the faster growing species that allows metabolically useful but slower growing species to remain within a community, and thereby the metabolic repertoire of the community will be maintained [2]. In a similar manner, antibiotic producing (ABP) strains can allow the growth of sensitive (ABS) species present in the same community provided they turn on antibiotic

production at higher signal levels. As a result, two species producing antibiotics against each other can stably coexist within the same community. Consequently, such a community will produce a variety of antibiotics that makes it difficult for external species to colonize the community.

Acknowledgments

Funded in part by Hungarian government grants OTKA 120650, EFOP-3.6.2-16-2017-00013 (supported by the European Union and co-financed by the European Social Fund) and ED_17-1-2017-0009 (within the National Bionics Program)

References

- [1] J. Juhász, Kertész-Farkas, D. Szabó, and S. Pongor. Emergence of collective territorial defense in bacterial communities: Horizontal gene transfer can stabilize microbiomes. *PlosOne*. 2014, 9, e95511.
- [2] J. Juhász, D. Bihary, A. Jády, S. Pongor and B. Ligeti B. Differential signal sensitivities can contribute to the stability of multispecies bacterial communities. *Biology Direct*. 2017, 22, 10.

POSTER SESSION

Surveying bacteriophages of the human gut

Balázs Ligeti^{1,2*}, Hőnich Ágnes Kata¹, János Juhász^{1,3} and Sándor Pongor¹

¹Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary;

²Institute of Medical Microbiology, Semmelweis University, Budapest Hungary

³In Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: ligeti.balazs@itk.ppke.hu

Keywords: microbiome, bacteriophage, gut microflora

Bacteriophages are one of the most abundant organisms in nature. Since the host organisms, bacteria, are widely considered as largest metabolic players in nature, bacteriophages themselves doubtlessly have an important ecological importance as they can alter, destroy, or else even stabilize bacterial communities. Bacteriophages also serve as a great gene reservoir and play a major role in horizontal gene transfer. The systematic study of bacteriophages is problematic however as they have small and highly divergent genomes. One of the many potential usages of bacteriophages is phage therapy - treating bacterial infections. Clinical trial studies have also investigated how phages can be used for altering metagenomes, or influencing, even reconstructing a microbiome. Despite these various interesting features and possibilities, large-scale systematic exploration of phages is still incomplete.

The aim of the present study is to develop bioinformatics methods capable of detecting phages in next generation sequencing (NGS) datasets. The method we plan to use for NGS is similar to the metagenomics binning algorithm developed previously [1]. In this algorithm, the reads of an NGS datasets are compared with a comprehensive and phylogenetically annotated database of phage sequences and the matching reads are mapped to their lowest common ancestor within the hierarchy. This will result in a ranked list of the most common phages or phage groups found within the NGS dataset. Key to this approach is the availability of a well-annotated phage sequence database. Preliminary experiments show that metagenomics datasets are often dominated by a single or a few phages.

Acknowledgments

Funded in part by Hungarian government grants OTKA 120650, EFOP-3.6.2-16-2017-00013 (supported by the European Union and co-financed by the European Social Fund) and ED_17-1-2017-0009 (within the National Bionics Program)

References

[1] L.S. Pongor, R. Vera, and Ligeti B. Fast and sensitive alignment of microbial whole genome sequencing reads to large sequence datasets on a desktop PC: application to metagenomic datasets and pathogen identification. PLoS One. 2014, 9, e103441.

POSTER SESSION

Role of cell-cell interactions between fluorescently labeled strains of S. cerevisiae in yeast colony formation

Tünde Gaizer¹, Báborka Pillér¹, Donát Pesti¹, Bence Makove¹, János, Juhász^{1,4}, Irene Stefanini², Bálint Péterfia¹, Attila Csikász-Nagy^{1,3}

¹Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, Budapest, Hungary

²University of Warwick, Coventry, UK

³Randall Centre for Cell and Molecular Biophysics, King's College London, London, UK

⁴3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

Traditional laboratory investigations of microbial colony formation are mostly focusing on single strains, however in their natural settings multiple, often genetically closely related strains live simultaneously. These mixed colonies provide opportunity to a great variety of interactions to occur. Besides the competition for nutrients, microbes can develop different strategies for better survival, like cooperation or competition. To better understand these kinds of interactions and their molecular background, we examine these interactions between the different strains of the budding yeast *Saccharomyces cerevisiae* to determine how they influence the future of mixed colonies.

We are also developing a systems biological approach to investigate these interactions. A high-throughput methodology is being developed to provide experimental data to be incorporated into computational models. Colony growth in liquid and solid media is measured to explore the differences between local and global interactions. Kinetic and agent-based models are developed to model growth of single and mixed cultures. Preliminary results of these studies are presented.

Acknowledgement:

The research has been supported by the European Union, co-financed by the European Social Fund(EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

Extended Partitioning Filter: Noise Filtering Machine Learning Method for Unbalanced Classes

Kristóf Szalóki^{1*}, Gábor Szűcs²

¹Budapest University of Technology and Economics

²Budapest University of Technology and Economics

*E-mail: szalokikristof@gmail.com

Keywords: noise filter, decision tree, unbalanced dataset, Extended Partitioning Filter, classification

Nowadays, the decision supporting systems based on the artificial intelligence are used worldwide in the analysis of dynamical systems in different areas like economy, medicine, biology. The systems need suitable amount and quality of data to work properly. However, machine learning methods, are often not prepared to handle the real life's inaccurate data. The aim of our paper was to create a noise filtering method, which can eliminate the noisy data meanwhile it keeps the outlier elements, even if the dataset is highly unbalanced. Based on Partitioning Filter [1] a new method, so called Extended Partitioning Filter (EPF) is suggested in this paper for these purposes.

Extended Partitioning Filter

The components of the method are as follows:

- **Partitioning**
The first component is partitioning, whereby the elements of the available data set are randomly divided into equal sized subsets.
- **Selection of good rules**
Using the elements of a subset, the method builds a decision tree from where it extracts its rule set. From this rule set the good rule set is defined based on the Agreement rate and Coverage. The size of the good rule set is a configurable parameter; and this selection component of EPF is a bioinspired phase of the method (similar to selection in the biological development).
- **Rule application**
The rules determine which class of all the other subsets will belong to. There are three possible outcomes of the step: (1) If the rule matches the element and correctly finds its class, then set a flag for 0, which represents the actual case and the root subset of the good rule set. (2) If the rule matches the element, but incorrectly defines its class, then set the flag to 1. (3) If a rule does not fit any case, then it does not change the flag.
- **Filtering**
Based on the flag attributes of the data set the method can determine that the elements are considered to be clear or noisy. Those elements whose sum of the flags is larger than a predefined noise threshold are eliminated.
- **Evaluation**
At the end of the iteration, there are three sets of data: the mostly clear, the clear, and the noisy. The part of the clear dataset is pushed backed to the mostly clear dataset as a base for the further iterations. The method terminates if the ratio of the actual found noise is less than a small proportion of the noises were found in the previous iteration.

Handling unbalanced datasets

In the case of noise filtering methods, it is especially important to handle unbalanced data. In our EPF method we solved it by an additional extra module: during over-sampling [2] elements of a class in a minority are enhanced by the SMOTE (Synthetic Minority Oversampling Technique) [3] technique. The whole process of Extended Partitioning Filter can be seen in Figure 1.

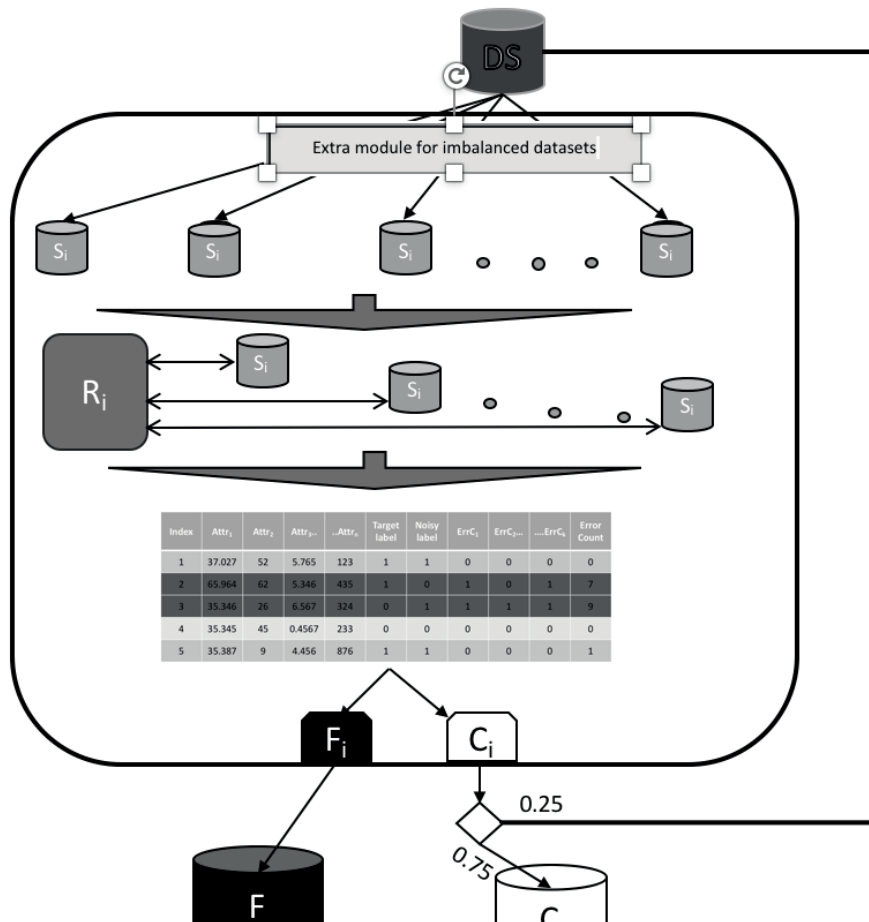


Figure 1. Process of Extended Partitioning Filter

Evaluation

The results of measurements clearly demonstrate the benefits of using EPF for noisy datasets, which is confirmed by the following diagram (as can be seen in Figure 2.). It shows the relative loss of accuracy [4] of the system with the EPF-filtered learning dataset ("Clear dataset" in the figure) compared to the noisy dataset (without EPF) as function of the noise level (the ratio of noisy element in the dataset); and it can be seen that the loss is always less by our suggested EPF.

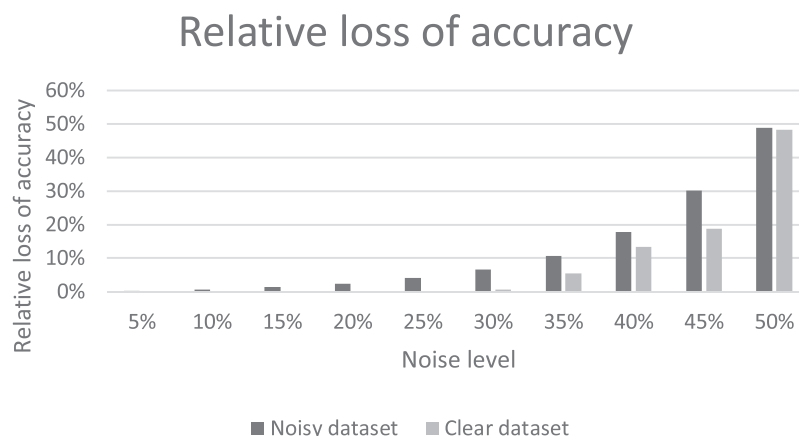


Figure 2. Relative loss of accuracy without and with EPF

Acknowledgments

The research presented in this paper has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013).

References

- [1] Zhu, X., Wu, X., & Chen, Q. (2003). Eliminating class noise in large datasets. In Proceedings of the 20th International Conference on Machine Learning (ICML-03), pp. 920-927.
- [2] Zheng, Z., Cai, Y., & Li, Y. (2015). Oversampling method for imbalanced classification. *Computing and Informatics*, 34(5), pp. 1017-1037.
- [3] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, pp. 321-357.
- [4] Sáez, J. A., Luengo, J., & Herrera, F. (2016). Evaluating the classifier behavior with noisy data considering performance and robustness: the equalized loss of accuracy measure. *Neurocomputing*, 176, pp. 26-35. doi: 10.1016/j.neucom.2014.11.086.

POSTER SESSION

Kinetic realizations of delayed dynamical models

Mihály Vághy^{1*}, Gábor Szederkényi¹, Gergely Szlobodnyik¹

¹Pázmány Péter Catholic University

²3in Research Group, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Esztergom, Hungary

*E-mail: vaghy.mihaly.andras@hallgato.ppke.hu

Keywords: chemical reaction networks, mass-action kinetics, delayed kinetic systems, linear programming

The nonnegative system classes are important tools in modeling such space/time varying phenomena, where the state variables are naturally positive (e.g. biochemical processes, reaction networks, compartmental systems). In a control system view the kinetic nonnegative systems are so-called universal descriptors [1]. In engineering practice it is often necessary that the delays appearing in the operation of processes (e.g. signal propagation, material flow) are modeled appropriately in order to describe the dynamic behavior precisely [2].

This work presents new results on the kinetic realization of delayed polynomial dynamical systems. A necessary and sufficient condition for the realizability of such models is given. Moreover, an algorithm is proposed for computing the so-called canonical realization of delayed models. An optimization based framework is generalized for computing dense linearly conjugate realizations, and all of the linearly conjugate realizations of kinetic systems using a fixed set of complexes. To prove the correctness of the algorithms, it is shown that dense linearly conjugate delayed reaction network realizations form a unique superstructure among all linearly conjugate realizations. Examples from the literature are presented to illustrate the applicability of the algorithms.

Acknowledgments

Supported by the ÚNKP-18-1-I-PPKE-47 New National Excellence Program of the Ministry of Human Capacities. The research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] P. Érdi and J. Tóth, *Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models*. Manchester, Princeton: Manchester University Press, Princeton University Press, 1989.
- [2] M. Roussel, "The use of delay differential equations in chemical kinetics," *The journal of physical chemistry*, vol. 100, no. 96, pp. 8323–8330, 1996.

POSTER SESSION

Towards Optimizing Hierarchical Temporal Memory for Parallel Computation

Csongor Pilinszki-Nagy^{1*}, Bálint Gyires-Tóth¹

¹ Dept. of Telecommunications and Media Informatics

Budapest University of Technology and Economics

Magyar Tudosok krt. 2, Budapest, Hungary, H-1117

*E-mail: csongor.pilinszkinagy@gmail.com

Keywords: neural network, deep learning, sequence learning, Hierarchical Temporal Memory, neocortex, optimization, sparse matrix, Sparse Distributed Representation, Spatial Pooler, Temporal Memory

Hierarchical Temporal Memory (HTM) is a machine learning method that utilizes the theory of how the neocortex functions. This theory presents a novel approach as to what it means to be intelligent and what are the core principles of this part of the human brain. It suggests that the neocortex is a homogeneous system arranged into a hierarchy of different abstraction layers, which do the same time-based pattern learning. Only one section of this hierarchy needs to be reverse engineered to understand and reimplement a human-like general intelligence. The HTM can be considered as a novel type of neural networks; it is suited to learn from sequential input, without providing the whole history of inputs to the network.

There has not been as much research on this topic as on deep learning [2], thus, it could also be a powerful machine learning method. The most promising results came from Numenta group and their book; the Biological and Machine Intelligence [1], which described their progress and possible implementations extensively. In this work possible underlying mechanics of the neocortex, arranged into four distinct tasks are introduced.

The network consists of four layers: an Encoder, a Spatial Pooler, a Temporal Memory, and a Decoder. The neurons in the network are arranged in columns. The columns are connected to the input and the cells inside the columns are connected to other cells in the network. The roles of the layers are specialized, the Encoder takes any type of data and turns it into a Sparse Distributed Representation. An SDR is a sparse binary representation of the input data, which retains similarity by overlapping bits for similar input data. The Spatial Pooler has two roles, it normalizes the input SDR while also learning spatial feature from the input data. The normalization means only a fixed number of columns in the network can be active and in a way that also retains similarity. The Temporal Memory is capable of learning time-based patterns by activating different cells in the active columns. The different active cells represent different context for the same input, this way what the network receives as input and what it received before is separable. The last layer decodes the HTM networks binary state to the next predicted value.

Numenta group released a proof of concept implementation, called NuPic, in Python, which demonstrates the learning capabilities of this type of network. At the time of writing optimized implementation of this theory does not exist at the time of writing.

The goal of our research is a complex network structure, which requires more computation to express as parallel operations. Every layer needs to have a parallel implementation to enable a faster execution on the suitable hardware. For this, the network is described as matrix and tensor operations. The challenge with

matrix representations is that the network consists of sparse activations and connections, thus, only a small fraction of its values are different from zero. Such a network can be efficiently described with sparse matrices and tensors. This type of sparseness requires novel sparse operators, like most of them are not optimized in linear algebra and machine learning libraries.

Our approach is to use the provided sparse matrix implementations and extend the dimensionality to 3 or more tensor dimensions, thus, enabling the Temporal Memory layer to be implemented using parallel operations. Furthermore, no HTM implementation uses massively parallel hardware, like Graphics Processing Units (GPUs), which massively support the current developments of machine learning. Using these methods combined with the matrix and tensor operations could multiple the speed at which the HTM network can efficiently process a larger amount of data. We expect that the modeling performance of HTM could be significantly improved this way. The HTM network already proved its learning capabilities on simple multidimensional scalar input sequences, in comparison with an LSTM baseline network, requiring fewer training epochs. The goal is to extend these capabilities to more complex input data sequences and reach inference speeds that enable real-time predictions.

Acknowledgments

The research in this paper has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013), by the BME-Artificial Intelligence FIKP grant of Ministry of Human Resources (BME FIKP-MI/SC), by Doctoral Research Scholarship of Ministry of Human Resources (ÚNKP-18-4-BME-394) in the scope of New National Excellence Program, by János Bolyai Research Scholarship of the Hungarian Academy of Sciences. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

References

- [1] Hawkins, J. et al. 2016. Biological and Machine Intelligence. Release 0.4. Accessed at <https://numenta.com/resources/biological-and-machine-intelligence/>.
- [2] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436.

REGISTERED PARTICIPANTS LIST

Bánhidy Anita Éva	banhidy.anita.eva@itk.ppke.hu
Bánhidy András	andras.banhidy@gmail.com
Battyányi Péter	battyanyi.peter@inf.unideb.hu
Becker Dániel Balázs	becker.daniel.balazs@hallgato.ppke.hu
Bellus Eszter	bellus.eszter@itk.ppke.hu
Búza Krisztián	buza@biointelligence.hu
Caesar Bálint	caesar.balint.gabor@itk.ppke.hu
Cardelli, Luca	luca.a.cardelli@gmail.com
Chakravarty, Suchana	suchanachakravarty@gmail.com
Ciliberto, Andrea	andrea.ciliberto@ifom.eu
Csabai Domonkos	csadom@gmail.com
Csány Gergely	csany.gergely@itk.ppke.hu
Csercsik Dávid	csercsik@itk.ppke.hu
Csikász- Nagy Attila	csikasz-nagy.attila@itk.ppke.hu
Csuhaj Varjú Erzsébet	csuhaj@inf.elte.hu
Dalchau, Neil	ndalchau@microsoft.com
Drexler Dániel	drexler.daniel@gmail.com
Dúzs Brigitta	duzsbrigitta@gmail.com
Érdi Péter	Peter.Erdi@kzoo.edu
Füzesi Krisztián	fuzesi.krisztian@itk.ppke.hu
Gaizer Tünde	tunde.gaizer@gmail.com
Gáspári Zoltán	gaspari.zoltan@itk.ppke.hu
Gönczy Tamás	tamas.gonczy@gmail.com
Gyöngy Miklós	gyongy.dermus@itk.ppke.hu
Harmat Zita	harmat.zita@itk.ppke.hu
Hartdégen Márton	hartdegen.marton@itk.ppke.hu
Hatvani Janka	hatvani.janka@itk.ppke.hu
Hegedűs István	ibe@itk.ppke.hu
Hinsenkamp Anett	hinsenkamp.anett@itk.ppke.hu
Iván Kristóf	ivan.kristof@itk.ppke.hu
Jensen, Lars Juhl	lars.juhl.jensen@gmail.com
Juhász János	juhasz.janos@itk.ppke.hu
Kaposi Ambrus	akaposi@inf.elte.hu
Keömley-Horváth Bence	horvathb94@gmail.com
Kerepesi Csaba	kerepesi@sztaki.hu
Kiss Rita Szilvia	kissrita.szilvia@gmail.com
Kiss-Tóth Annamária	kiss-toth.annamaria@itk.ppke.hu
Kistóth Éva Mercédesz	kistoth.eva.mercedesz@hallgato.ppke.hu
Kosztá Imola	koszta@inf.elte.hu
Kovács Levente	kovacs@uni-obuda.hu
Kozsik Tamás	Tamas.Kozsik@elte.hu

Kun Péter	kunpter@gmail.com
Kwiatkowska , Marta	marta.kwiatkowska@cs.ox.ac.uk
Ladics Tamás	tamas.ladics@gmail.com
Laki András	laki.andras.jozsef@itk.ppke.hu
Laurenti, Luca	luca.laurenti@cs.ox.ac.uk
Lente Gábor	lente@gamma.ttk.pte.hu
Ligeti Balázs	ligeti.balazs@itk.ppke.hu
Marosán Péter	marosan.peter@itk.ppke.hu
Menche, Jörg	JMenche@cemm.oeaw.ac.at
Nagy Olivér	nagy.oliver@itk.ppke.hu
Németh Erik	erik.nemeth96@gmail.com
Oláh András	olah.andras@itk.ppke.hu
Orosz Áron	emegtap@gmail.com
Otero-Muras, Irene	ireneotero@iim.csic.es
Ősz Katalin	oszkgamma.ttk.pte.hu
Paccanaro, Alberto	Alberto.Paccanaro@rhul.ac.uk
Pesti Donát	hondonop@gmail.com
Pilinszki-Nagy Csongor	csongor.pilinszkinagy@gmail.com
Pillér Biborka	pillerborka@gmail.com
Pongor Csaba Sandor	csaba.pongor@gmail.com
Pongor Sandor	pongor@icgeb.org
Przulj, Natasa	natasa@cs.ucl.ac.uk
Reguly István	reguly.istvan.zoltan@itk.ppke.hu
Répai Attila	repai.attila@hallgato.ppke.hu
Schaffer Katalin	sch.katalin17@gmail.com
Sváb Gergely	gszab92@gmail.com
Szabó, Zoltán Csaba	zoltantaylor3@gmail.com
Szalóki Kristóf	szalokikristof@gmail.com
Szederkényi Gábor	szederkenyi@itk.ppke.hu
Szep, Gregory	gregory.szep@kcl.ac.uk
Szlobodnyik Gergely	szlger91@gmail.com
Szmolyan, Peter	szmolyan@tuwien.ac.at
Szolgay Péter	szolgay.peter@itk.ppke.hu
Szűcs Gábor	szucs@tmit.bme.hu
Tasi Benedek József	tasibenedekjozsef@gmail.com
Tihanyi Attila	tihanyi.attila@itk.ppke.hu
Tóth János	jtoth@math.bme.hu
Ursula Lenhert	ursula.lenhert@mckinsey.com
Vághy Mihály	vaghy.mihaly.andras@hallgato.ppke.hu
Vaszil György	vaszil.gyorgy@inf.unideb.hu

Contents

Programme of the Conferences	5
BIOINSPIRED ANALYSIS OF DYNAMICAL NETWORKS	6
Keynote, Invited lecture	7
Dynamical systems perspective in neuroscience - historical and current approaches	7
Computing Reliably with Molekular Walkers	8
Model-based design of synthetic biological pattern formation	9
Engineering complexity in biochemical networks: methods and tools for the automated design of biochemical bistable switches and limit cycle oscillators	10
Automatic model selection, automatic initial estimates	11
Selected talks	13
Continuous Blood Pressure Measurement with Camera-based Optical Methods	13
Programming by Multiset Transformations and the Membrane	16
Computational Paradigm	16
Multiple time-scale dynamical systems in cell biology	20
On Dynamics of Networks of Watson-Crick DOL Systems	22
Spatio-temporal Modelling of Tumor Progression and Angiogenesis	25
Chemical analogues of biological periodicity: the immanent potential of pH oscillators	29
Design and Development of an Anatomically Correct, Biomechatronic Hand Model	30
Chemical reaction models of gene expression boundary	36
dynamics in synthetic bacterial colonies	36
Molecular Filters for Noise Reduction	38
PROTEIN NETWORK ANALYSIS	40
Keynote, Invited lecture	41
Morphisms of chemical reaction networks	41
Network analysis and visualization with STRING and Cytoscape	42
From protein-protein to human-machine interactions	43
Mining the Integrated Connectedness of Biomedical Systems	44
Answering questions in biology and medicine by making inferences on networks	45
Selected talks	46
MiStImm: an agent-based simulation tool to study the self-nonsel	46
discrimination of the adaptive immune response	46
Integrating molecular databases in a proteome-wide	49
simulation environment.	49
Exploring inhibitory logics in protein-protein interaction networks	52
Tumor growth modeling using formal reaction kinetics analogy	53
A Bayesian Ranking Approach for Drug-Target Interaction	56
Prediction and a Web Application Containing the Predictions	56
Binding sites in disordered regions might contribute to the complexity of the postsynaptic protein network	58
Comparison of the Deterministic and Stochastic Approaches to	60

Michaelis-Menten Kinetics	60
Ensemble-based interpretation of the dynamic allostery in a PDZ domain	63
Reachability analysis of discrete state chemical reaction networks obeying conservation laws	65
POSTER SESSION for the two Conference	68
Effect of protein aggregation on protein complex assemblies	69
Dynamical modeling of the most important pathways in cellular metabolism	71
Towards a systematic census of bacterial quorum sensing genes in public databases	72
Network-based prediction of drug-combination efficiency based on expression data on off target genes	75
Modelling genomic and metabolic responses of microbial communities to antibiotics	77
Surveying bacteriophages of the human gut	79
Role of cell-cell interactions between fluorescently labeled strains of <i>S. cerevisiae</i> in yeast colony formation	80
Extended Partitioning Filter: Noise Filtering Machine Learning Method for Unbalanced Classes	81
Kinetic realizations of delayed dynamical models	84
Towards Optimizing Hierarchical Temporal Memory for Parallel Computation	85
REGISTERED PARTICIPANTS LIST	87

