Sixth International Conference in Code Biology

Friedrichsdorf (Germany) 3-7 June 2019

ABSTRACTS

1 Ádám Kun
2 Ádám Radványi
3 Almo Farina
4 Andjela Rodić
5 Anna Aragno
6 Andres Kurismaa
7 Branko Dragovich
8 Candice Shelby
9 Catherine McCusker
10 Charles Carter
11 Christian Michel
12 David Ellerman
13 Diego Gonzalez
14 Elena Fimmel
15 Hanna Schumacher
16 Jacques Demongeot
17 Kirti Prakash
18 Ludmila Lacková
19 Małgorzata Wnętrzak
20 Marcella Faria
21 Marcello Barbieri
22 Markus Gumbel
23 Markus Schmidt
24 Martin Starman
25 Morten Tønnessen
26 Nadir Maraldi
27 Nataša Mišić
28 Nediljko Budisa
29 Nikola Štambuk
30 Omar Paredes
31 Paul Sorba
32 Paweł Błażej
33 Pedro Marijuán
34 Peter Dittrich
35 Peter Wills
36 Robert Prinz
37 Rasmus Gahrn-Andersen
38 Sergey Petoukhov
39 Simone Giannerini
40 Stephen Cowley
41 Wanderley dos Santos
42 Willem Beekman

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Expansion scenarios of the amino acid repertoire suggested by ancestral protein elements

Ádám Kun [a,b] and Ádám Radványi [c]
[a)] Theoretical Biology and Evolutionary Ecology Research Group, Budapest, Hungary
[b)] Parmenides Centre for the Conceptual Foundation of Science, Munich, Germany
[c)] Department of Plant Systematics, Ecology and Theoretical Biology,
Eötvös University, Budapest, Hungary.
kunadam@elte.hu

The origin of translation is among the most intractable puzzles in the origin of life, since most traces are washed away by billions of years. However, in recent years, several reports revealed the existence of promiscuous loop-like protein structures. These sets are grouping sequences with low sequence identity, yet their structures and functions show high similarities. Furthermore, the functions of these elements are mostly associated with nucleotides, nucleic acids and binding of small substrates. Both their promiscuity, their short length of 20–30 amino acids, and functions associated with key metabolic processes indicates that these specimens might be connected with the first-ever appearance of coded protein structures. A logical interest is to make a closer observation on these structures, as they would contain further information regarding the archaic translation machinery. We were interested whether these protein elements can be used to infer a simplified translation process, and to establish possible scenarios for the expansion to the present canonical 20 amino acids. First, we calculated the mutual information contained in the individual protein element sets, based on a method previously devised by Yockey. This method is particularly helpful in predicting those amino acids which are not present in a given position, yet might occupy it in the future or in the past. We also assayed for those scenarios for amino acid repertoire expansions, which would be the most efficient way to make all known promiscuous protein element attainable. Our analysis shows that these structural groups have low information content, making an equivocal suggestion: Either the first protein elements consisted of a few amino acids, or these structures were produced via a less reliable translation process resulting in statistical polymers having high tolerance for errors (c.f. Barbieri's proposal for the evolution of translation and the genetic code). The repertoire expansion scenarios also reveal that some amino acid might had a crucial role in accessing these ancestral structures. Connections with the operational genetic code are also discussed.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

On the origin of translation and the possible habitat of LUCA based on a simple neutral model of amino acid composition

Ádám Radványi[a] and Ádám Kun[b,c,d]

[a)]Department of Plant Systematics, Ecology and Theoretical Biology, Eötvös University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary. [b)] Evolutionary Systems Research Group, MTA, Centre for Ecological Research, Hungarian Academy of Sciences, Klebelsberg Kuno u. 3., 8237 Tihany, Hungary [c)] Parmenides Centre for the Conceptual Foundation of Science, Kirchplatz 1, D-82049 Munich/Pullach, Germany; [d)] MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary

adamradvanyi117@gmail.com

The origin of the genetic code is a notorious problem, however we also know that The Last Universal Common Ancestor (LUCA) already harboured this standard codon mapping. As coding theory suggests the understanding of codes via the distribution of coded symbols, the ancestral amino acid distribution of LUCA might reveal something about the genetic code. Unfortunately, present phylogenetic methods are still unable to reliably identify LUCA, yet existing data could allow us to make educated guesses regarding its nature. One possibility would be to make deductions based on the amino acid composition of present organisms. We suggest a simple Markov model based on the neutral theory of molecular evolution. Assuming that neutral mutations has the greatest impact on proteomes, derived transition matrices of neutral mutations (e.g. the PAM probability matrix) should determine the expected equilibrium frequency distribution on the long term. Exploiting this suggested evolutionary progress, the key aspect is to concentrate on how far a given organism is from the equilibrium. This is measured by information-theoretic divergence indices. Here we show that halophile archaeas and some thermophile prokaryote (the latter includes both Archaea and Bacteria) provide the highest distance from equilibrium. Although high divergences could be caused by strong environmental selection on the amino acid composition, Dirichlet regression models imply that only halophile composition might be the result of selection. Contrary to previous works, similar selection trends were not apparent in proteomes related to high temperature. In conclusion, thermophiles might have preserved crucial information about the LUCA, which is also supported by their basal position in the Tree of Life. Further implications of the suspected ancestral amino acid compositions, regarding the origin and diversification of life, and its possible connections with the origin of the genetic code are also discussed.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Acoustic codes: A multiscale approach

Almo Farina
Department of Pure and Applied Sciences
Urbino University, Urbino, Italy
almo.farina@uniurb.it

Acoustic events may be considered as units of intra or interspecific communication and are composed of biophonies often associated to geophonies and technophonies. The combination of these three components produces a complex information system that may be utilized by individual species to intercept a broad spectrum of resources. The encoding process applied to the ecoacoustic events that produces the acoustic codes has been recently refined by Farina et al. (2017) using the Ecoacoustic Event Detection and Identification model (EEDI). This model generates three digit codes that in turn are expression of the Acoustic Complexity Index metrics. EEDI procedure is sensitive to temporal resolution at which the acoustic files are processed. To reduce the subjectivity in the choice of an appropriate scale a multiscale approach has been utilized in the encoding process. For this, ten temporal scales (1,2,4,8,16,30,60,120,240,360 second) are proposed. Regressing the log of scale and the log of the number of Ecoacoustic Events a fractal dimension *d* is obtained. *d* seems a good indicator of the distribution of acoustic information that circulate inside acoustic files and that species utilize to communicate and to track resources.

Reference

Farina A., Gage S.H. and Salutari P. 2018. Testing the Ecoacoustics Event Detection and Identification (EEDI) Model on Mediterranean Soundscapes. *Ecol. Indicators*, 85: 698-715.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Possible code in dynamical properties of diverse bacterial immune systems

Andjela Rodić [1,2], Bojana Blagojevic[3], Magdalena Djordjevic[3], Marko Djordjevic[1]
[1] Faculty of Biology, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia
[2] Multidisciplinary PhD program in Biophysics, University of Belgrade,
[3] Institute of Physics, University of Belgrade, Pregrevica 118, 11000 Belgrade, Serbia
andjela.rodic@bio.bg.ac.rs

A major goal of quantitative systems biology is to discover design principles of biological systems. Such design principles are common properties of otherwise mechanistically diverse systems, which can be interpreted as code in their functioning. We here investigate design principles (code) of different bacterial immune systems, in particular their common dynamical properties that allow efficient protection against incoming viruses, while at the same time preventing autoimmunity (destruction of its host). For this, we use a combination of statistical thermodynamics modeling of system transcription regulation, and dynamical modeling of generation of relevant molecular species (RNA and proteins). In particular, we look at restriction-modification (R-M) and CRISPR-Cas systems, which both defend host cell from foreign DNA. R-M are more rudimentary systems in which restriction endonuclease cuts specific DNA sequences, while methyltransferase protects the same sequences from cutting. Once the plasmid carrying the system enters a new host cell, it is essential for the cell survival that expression of system enzymes is tightly controlled, so that the host genome is not destroyed. R-M systems use various regulatory features for such control - have different architectures (convergent or divergent gene organization), encode a specialized control (C) protein, involve overlapping promoters, inefficient translation of some of its transcripts etc.
The same defense function is fulfilled in a more sophisticated way by the advanced system, CRISPR-Cas, which encodes small interfering RNAs (crRNAs), which recognize invasive DNA targets, allowing their consequent destruction. CRISPR-Cas in *E. coli* is normally silent, and two main features of CRISPR-Cas activation are its highly cooperative transcription control, and fast non-specific degradation of long CRISPR transcripts. By extensively computationally perturbing these key regulatory features in representative R-M systems [1-3], and in CRISPR-Cas [4-5], we show that these diverse features can be explained in terms of few simple common dynamical properties, most importantly in terms of fast transition from OFF to ON state (promoting efficient defense against viruses), and a delayed expression of toxic molecules (preventing autoimmunity).

[1] Rodic, A., Blagojevic, B., Zdobnov, E., Djordjevic, M., Djordjevic, M. (2017) Understanding key features of bacterial restriction-modification systems through quantitative modeling. *BMC systems biology* **11**:377.
[2] Morozova, N., Sabantsev, A., Bogdanova, E., Fedorova, Y., Maikova, A., Vedyaykin, A., Rodic, A., Djordjevic, M., Khodorkovskii, M., Severinov, K. (2015) Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. *Nucleic acids research* **44**:790.
[3] Bogdanova, E., Djordjevic, M., Papapanagiotou, I., Heyduk, T., Kneale, G., Severinov, K. (2008) Transcription regulation of the type II restriction-modification system AhdI. *Nucleic acids research* **36**:1429.
[4] Rodic, A., Blagojevic, B., Djordjevic, M., Severinov, K., Djordjevic, M. (2017) Features of CRISPR-Cas Regulation Key to Highly Efficient and Temporally-Specific crRNA Production. *Frontiers in microbiology* **8**:2139.
[5] Djordjevic, M., Djordjevic, M., Severinov, K. (2012) CRISPR transcript processing: a mechanism for generating a large number of small interfering RNAs. *Biology direct* **7**:24.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

With Body in Mind – Embodied Language

Anna Aragno
National Association for the Advancement of Psychoanalysis
140 West End Avenue, New York, NY, 10023, USA
annaragno@earthlink.net

….for in the psychical field, the biological field does in fact play the part of the underlying bedrock
S. Freud, 1937, p. 252

Freud established Psychoanalysis as an experiential methodology to uncover unconscious processes and phenomena and decode all non-conscious forms of meaning-making. While it was the analysis of the pictographic dream, as prime specimen, that gave Freud access to the structure and forms of the unconscious, it is through a linguistic code that we interpret its meanings. This presentation addresses the psychoanalytic interpretive system which, while constrained by language, is not fundamentally linguistic, and will be in two parts: first I will discuss a paradigm suited to the bio-semiotic study of all forms of human inter-action and communication. Second, I will address some non-linguistic functions to which our cultural code is put, specifically regarding what *is* and what *can* be done with words.

All knowledge is filtered through and pinned down by some semiotic system; within the limitations and the potential of language, we crystallize new concepts through its nominal capacities, by putting things into words. When listening to the unconscious, however, psychoanalysts adopt a highly specialized attentional stance employing all the senses and emotions as instruments of attunement.

Psychoanalysis is a method created by language: words are our instruments and medium. This positions us optimally to observe different semiotic levels, from organic symptoms and emotional signals, to the sign and symbolic forms of language. Each of these expose different levels of organismic organization best systemized in a *developmental continuum* (Aragno 1997/2016) beginning in pre- and proto-semiotic forms and culminating in symbolic language-use. This multiple-code model reframes theoretical understanding around epigenetic and morphological principles suitable for developmental processes bridging biological and psychological organizations. From this revised and broadened meta-theoretical base, I will illustrate how language absorbs deeper unconscious impulses, presenting ten to which speech may be put, ending with two that are specific to a dialogue designated as the “talking cure.”

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Towards a code model of neural afferent synthesis

Andres Kurismaa
Department of Philosophy and History of Science
Charles University, Czechia
andres.kurismaa@gmail.com

A particularly interesting, if challenging question for qualitative accounts of biological information has been the problem of neural signal integration, including its possible anticipatory and autonomous aspects. Here, it is considered how new types of experiments on the mechanisms underlying action potential generation in single neurons (Sardi et al 2017) may be particularly relevant in this case, and complement earlier studies on the mechanisms of afferent integration and anticipation in single brain cells (Anokhin 1984). Although in different aspects, it is argued that both of these analyses reveal unexpected complexity and selectivity in the basic processes of signal "summation" in single cells, and possibly, allow to consider these key neural processes from a code biological perspective.

Thus, based on early evidence from electron microscopy, as well as pharmacological and physiological studies, P.K. Anokhin and colleagues (Anokhin 1984) suggested that the biophysical propagation and summation models of signal transmission (prevalent at the time, and in many aspects even now) may contain serious fallacies when applied from the effector side of a neuron (axon) to its receptor side (dendrites and soma), as doing so would obscure the whole problem of information specificity and transmission – how a particular set of afferent inputs can be mapped to correspondingly specific outputs in individual neurons and functional systems? Furthermore, compatible experimental evidence has recently emerged revealing new types of integrative and selective processes at the basis of spike generation in single cells (Sardi et al 2017). By showing evidence conflicting with the notion of isotropic summation, as well as demonstrating non-occurence of simple summation and subtraction effects in combined intra- and extracellular stimulations (Sardi et al 2017), these novel findings may call for current biophysical models of spike generation to be significantly revised – with a view on possible multiple independent threshold elements within each cell that are anisotropically activated (Sardi 2017), and in line with functional systems theory, may express differential transmitter sensitivity, nuclear pathways and metabolism (Anokhin 1984). If supported by further evidence, these could be important steps towards revealing the code complexity of afferent synthesis in single brain cells, and more generally, understanding the possibility of qualitative distinctions being made and retained at the level of elementary neural processing (Tse 2013).

References:

Anokhin, P. K. (1984). Systems analysis of the integrative activity of the neuron. *The Pavlovian Journal of Biological Science*, *19*(2), 43-101.

Sardi, S., Vardi, R., Sheinin, A., Goldental, A., & Kanter, I. (2017). New Types of Experiments Reveal that a Neuron Functions as Multiple Independent Threshold Units. *Scientific Reports*, *7*(1), 18036.

Tse, P. (2013). *The Neural Basis of Free Will: Criterial Causation*. Mit Press.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

p-Adic Properties of the Genetic Code

Branko Dragovich[1,2]
[1]Institute of Physics, University of Belgrade, Belgrade
[2]Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade, Serbia
dragovich@ipb.ac.rs

The genetic code is a connection between 64 codons and 20 canonical amino acids with one stop signal. It is almost unique in all living organisms, in spite of a huge number of mathematical possibilities. Structure of the vertebrate mitochondrial (VM) code is simpler than the standard one, which can be regarded as slight variation of the VM code. In the VM code, an amino acid is coded by one, two or three codon doublets. When two codons code the same amino acid, in the informational sense they are close (similar). It is shown (see e.g. [1,2,3]) that the p-adic distance is a simple mathematical tool for description of codon closeness (similarity).
p-Adic distance (p is a prime number) is the most useful example of ultrametrics and it is a a basic concept in p-adic analysis and its applications in modeling physical and biological systems with hierarchical structure, for a recent review see e.g. [4]. In p-adic approach to modeling genetic code, p-adic distance plays a central role. p-Adic distance between two integer numbers is related to divisibility of their difference by prime number p: larger divisibility -- smaller distance. When p-adic distance is smaller, the related numbers are p-adic closer (more similar).
In this approach to modeling the genetic code one assigns 64 natural numbers in the form $a_0+a_1\,5+a_2\,5^2$ to the 64 codons appropriately identifing nucleotides in codons with digits a_i in these numbers. I take the following identification: C(Cytosine) = 1, A(Adenine) =2, U(Uracil)= T(Thymine) =3, and G(Guanine) = 4. With respect to the smallest 5-adic distance between codons one obtains 16 quadruplets. Each of these quadruplets splits into two doublets under the smallest 2-adic distance. Then each of these 32 doublets codes one amino acid or stop signal in the VM code. This p-adic approach has been extended to some other aspects of the genetic code: similarity between amino acids, evolution of the genetc code, genetic code as a p-adic network, similarity between bioinformation sequences.
In this talk, I shall present main results and discuss a wobble base pairing from the p-adic point of view.

References

[1] B. Dragovich and A. Dragovich, ``A p-adic model of DNA sequence and genetic code'', p-Adic Numbers Ultrametric Anal. Appl. **1** (1), 34--41 (2009), arXiv:q-bio.GN/0607018v1.
[2] B. Dragovich and A. Dragovich, ``p-Adic modelling of the genome and the genetic code'', Computer J. **53** (4), 432--442 (2010), arXiv:0707.3043v1 [q-bio.OT].
[3] B. Dragovich, A. Yu. Khrennokov and N. Ž. Mišić, ``Ultrametrics in the genetic code and the genome'', Appl. Math. Comput. **309**, 350--358 (2017), arXiv:1704.04194 [q-bio.OT].
[4] B. Dragovich et al., ``p-Adic mathematical physics: the first 30 years'', p-Adic Numbers Ultrametric Anal. Appl. **9** (2), 87--121 (2017), arXiv:1705.04758 [math-ph].

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Biological Coding and Honor Killings

Candice L. Shelby
Professor of Philosophy, University of Colorado Denver
Campus Box 179, P. O. Box 173364
Denver, CO 80217-3364
Candice.Shelby@ucdenver.edu

It would seem impossible for most of us to even conceive of killing a sister, daughter, or mother, but in numerous regions in the world (and in sub-communities in most countries in western Europe and in the U.S.) it happens with troubling regularity. How is it that the (mostly men, but sometimes women, acting in their capacity as patriarchal stand-ins) in these communities can come to be so constituted that such a thing is not only possible, but is experienced as a necessary part of being a responsible member of a family? Philosopher Robert Paul Churchill's recent book argues that it requires, in the first instance, an Honor-Shame Culture of certain type, but in the second instance, the development of a particular type of violence-prone men and a particular experience of the sense of self. This paper argues that violence, control, and a very particular conception and embodiment of masculinity are encoded into these individuals' psyches in a way that allows them, contrary to all intuitions about evolution, to be willing and able to kill female family members, even their mothers. The encoding referred to here is not merely a matter of the codes involved in particular versions of the MAO-A, DRD2, or DAT-1 gene (although given the inter-group marriages often required in these societies, those versions may well be more prevalent than in the general population), or of social codes. This encoding is rather a matter of the expression of genetic codes, aberrant or not, within a complex dynamic system that is perturbed in specific embodied, embedded, and enactive ways. In the case of males reared within a patriarchal, competitive, and shame-controlled society, in which their sexuality is both lauded and encouraged, and, what is more, strongly contrasted with that of their sisters, mothers, and cousins, but in which simultaneously the honor of the entire family rests on the chastity and modesty of its females, the very embodied experience of love and sexuality is coded for violence.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Hierarchical pattern formation during amphibian limb regeneration

Catherine McCusker and Warren Vieirra
Department of Biology, University of Massachusetts, Boston, MA, USA
mccusker.catherine@gmail.com

In 1901 T.H. Morgan proposed in "Regeneration" that pattern formation in amphibian limb regenerate is a stepwise process. Since this time, biologist have continued to piece together the molecular components of this process to better understand the "patterning code" during limb regeneration. While a number of models have been proposed, two main camps of thought have emerged. One camp is invested in the idea, known as the "morphogen hypothesis", that patterning emerges from the localized expression of signaling molecules, which produce differing cellular responses depending on the intensity of the signal. The other camp hypothesizes that cells retain memory of their patterning information, and tap into this information in the cells of the remaining stump tissue to generate new cells with the missing pattern information, a process called intercalation. A growing body of evidence supports the possibility that these two mechanisms are not mutually exclusive. Here, we propose our theory of hierarchical pattern formation, which consists of 4 basic steps. The first is the existence of cells with positional memory. The second is the communication of positional information through cell-cell interactions in a developmental-permissive environment. The third step is the induction of molecular signaling centers. And the last step is the interpretation of these signals by specialized cell types.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Hierarchical groove discrimination by Class I and II aminoacyl-tRNA synthetases reveals a palimpsest of the operational RNA code in the tRNA acceptor-stem bases

Charles W. Carter, Jr[1] and Peter R. Wills[2]

[1] Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7260
carter@med.unc.edu
[2] Department of Physics, Centre for Computational Evolution, and Te Ao Marama, University of Auckland, PB 92109, Auckland 1142, New Zealand
p.wills@auckland.ac.nz

A new regression model identifies determinants for bacterial tRNAs cognate to Class I and II aaRS (1). Groove discrimination involves only bases 1, 2, and 73; is hierarchical; and suggests that new rules joined pre-existing ones as the coding alphabet grew. A thermodynamic rationale implies that Class-dependent aaRS secondary structures exploited differential tendencies to form the 3' terminal (Class I) hairpin, simultaneously differentiating amino acid types, aaRS, and cognate tRNAs; aaRSs evolution was thus necessary and sufficient to embed protein folding rules (2) into acceptor stem bases (3,4). Base 2 determines groove recognition in a manner consistent with the codon table, confirming ancestral codon-anticodon pairing in the acceptor stem and productive coding interaction of tRNA minihelices with proto-mRNA. The operational RNA code emerges as five overlapping patterns of acceptor stem bases distinguishing groups of amino acids (1).

1. Carter, C.W., Jr and Wills, P.R. (2018) NAR, In revision.
2. Carter, C.W., Jr. and Wolfenden, R. (2016) RNA Biology, 13:145
3. Wills, P.R. and Carter, C.W., Jr. (2018) BioSystems, 164, 155
4. Carter, C.W., Jr and Wills, P.R. (2018) MBE 35, 269

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Circular code motifs in genes

Christian J. Michel
Université de Strasbourg, CNRS
300 Boulevard Sébastien Brant, 67400 Illkirch, France
c.michel@unistra.fr

(joint work with Dila G., Nguefack Ngoune V., Poch O., Ripp R., Thompson J.D.)

The *X* motifs, i.e. motifs from the circular code *X* which has been found in genes of bacteria, archaea, eukaryotes, plasmids and viruses (Michel, 2017, 2015; Arquès and Michel, 1996), have the property to retrieve, maintain and synchronize the reading frame in genes. In a recent study of the *X* motifs in the complete genome of the yeast, *Saccharomyces cerevisiae*, it was shown that they are significantly enriched in the reading frame of the genes (protein-coding regions) of the genome (Michel et al., 2017). It was suggested that these *X* motifs may be evolutionary relics of a primitive code originally used for gene translation. In a large scale analysis involving complete genomes from four mammals and nine different yeast species, we highlight specific evolutionary pressures on the *X* motifs in the genes of all the genomes, and identify important new properties of *X* motif conservation at the level of the encoded amino acids (Dila et al., 2019). We then compare the occurrence of *X* motifs with existing experimental data concerning protein expression and protein production, and report a significant correlation between the number of *X* motifs in a gene and increased protein abundance. In a general way, this work suggests that motifs from circular codes may represent functional elements located within the coding regions of extant genomes.

References

Arquès D.G., Michel C.J. (1996). A complementary circular code in the protein coding genes. Journal of Theoretical Biology 182, 45-58.

Michel C.J. (2015). The maximal C3 self-complementary trinucleotide circular code X in genes of bacteria, eukaryotes, plasmids and viruses. Journal of Theoretical Biology 380, 156-177.

Michel C.J. (2017). The maximal C^3 self-complementary trinucleotide circular code *X* in genes of bacteria, archaea, eukaryotes, plasmids and viruses. **Life** 7, 20, 1-16.

Michel C.J., Nguefack Ngoune V., Poch O., Ripp R., Thompson J.D. (2017). Enrichment of circular code motifs in the genes of the yeast Saccharomyces cerevisiae. Life 7, 52, 1-20.

Dila G., Michel C.J., Poch O., Ripp R., Thompson J.D. (2019). Evolutionary conservation and functional implications of circular code motifs in eukaryotic genomes. Biosystems 175, 57-74.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Bateson was right: Information is differences

David Ellerman
University of California Riverside
University of Ljubljana, Slovenia
david@ellerman.org

Gregory Bateson is considered as a precursor ('avant le lettre') of biosemiotics (1). One of Bateson's seminal themes was the role of differences and particularly the idea of communicated information as a "difference that makes a difference." There has been a recent development in mathematical logic which gives rise to a new *logical* theory of information that proves Bateson was right.

Ordinarily, the most basic form of logic is the Boolean logic of subsets (a.k.a. 'propositional' logic), but that turns out to be only half the story. The notion of a subset has a category-theoretic dual in the notion of a partition (= quotient set = equivalence relation) so the 'other half' is the newly developed logic of partitions (2). The quantitative version of the Boolean logic of subsets is the logical (Laplace-Boole) notion of probability, and the corresponding quantitative version of partition logic is the logical notion of information (or logical entropy) (3). The *differences* or *distinctions* of a partition are the pairs of elements that are distinguished by the partition (i.e., are in different equivalence classes or blocks of the partition), and the *logical notion of information* in a partition is just the normalized number of differences given by the partition. This is the logical explication of Bateson's notion of information as differences (and thus the communication of information as the "differences that make a difference"). The logical theory of information is the foundational theory that subsumes the Shannon theory in the sense that all the Shannon definitions of simple, joint, conditional, and mutual entropy are obtained by a uniform requantifying transformation from the corresponding logical definitions.

The basic idea of the requantification is that the logical entropy is the (normalized) count of the number of distinctions while the Shannon entropy is the average number of binary partitions (or yes-or-no questions) needed to make the same distinctions, i.e., the average binary code size needed to uniquely designate or encode the distinguished entities or messages. Thus the Shannon theory is repositioned as the specialized requantification of logical theory for the purposes of coding and communication. Hence these developments serve to give the mathematical version of Bateson's notion of information-as-differences and then to rigorously connect it to the notions of coding and communication whose application to living systems is the subject matter of biosemiotics and code biology.

References

(1) Hoffmeyer, Jesper, ed. 2008. *A Legacy for Living Systems: Gregory Bateson as Precursor to Biosemiotics*. Springer.com

(2) Ellerman, David. 2010. "The Logic of Partitions: Introduction to the Dual of the Logic of Subsets." *Review of Symbolic Logic* 3 (2 June): 287–350. Available at: Ellerman.org

(3) Ellerman, David. 2017. "Logical Information Theory: New Foundations for Information Theory." *Logic Journal of the IGPL* 25 (5 Oct.): 806–35. https://doi.org/10.1093/jigpal/jzx022. Available at: Ellerman.org.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The Rumer's transformation: a symmetry puzzle standing half a century

Diego Gonzalez and Simone Giannerini
University of Bologna, Department of Statistical Sciences
CNR/IMM Bologna Section
gonzalez@bo.imm.cnr.it

In 1966, only a few months after the complete elucidation of the standard nuclear genetic code, the Russian theoretical physicist Yury Borisovich Rumer uncovered the existence of a particular symmetry: when the keto-amino, or Rumer's transformation, is applied to the bases of a codon, the degeneracy of the represented amino acid changes from 4, i.e. an amino acid family, to non-4, i.e. degeneracy 1, 2, or 3 (or viceversa). After half a century of such discovering and despite the universality of the Rumer's symmetry (is shared by the most representative versions of the genetic code), little is know about its origin and its possible biological significance. Here we show that the Rumer's symmetry has originated in an ancestral version of the genetic code, i.e., the *pre-early* code, implying codons of length 4 (4 bases per codon) and corresponding ancient adaptors. The Rumer's symmetry, thus, arises as a natural consequence of the stereo-chemical symmetries present in the ancestral synthesis machinery at the time of the pre-early code. Moreover, the conservation of the Rumer's symmetry through relevant evolutionary periods suggests a connection with key extant biological features. Intriguing possibilities in this sense are such of error detection/correction, control over the synthesis of proteins, and frame maintenance.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

A general approach to circular codes with a view to Biology

Elena Fimmel and Lutz H. Strüngmann
Institute for Mathematical Biology, Faculty of Computer Science,
Mannheim University of Applied Sciences, Paul-Wittsack-Str. 10
68163 Mannheim, Germany
e.fimmel@hs-mannheim.de
l.struengmann@hs-mannheim.de

(joint work with Christian Michel, Francois Pirot, and Jean-Sébastien Sereni)

Circular codes have been found in large populations of genes and are assumed to play an essential role in maintaining the reading frame during the translational process in the ribosome. Moreover, circular codes are also used in signal processing and other areas dealing with data transmission. In the former case the genetic alphabet {A,C,G,T} serves as the set of letters while in the latter case circular codes over general alphabets are of interest. This is our motivation to investigate l-letter circular codes over alphabets with an arbitrary number n of letters.
In this talk we present some recent results obtained in [1] on the structure, size and number of such circular codes over general alphabets. We will also consider the important subclasses of comma-free and strong comma-free codes which have stronger error detecting properties than circular codes when it comes to frame recognition. Several examples will be given to illustrate our findings and to show that the previous results on circular codes in the genetic context follow from our new findings. Finally, we put our work in a biological context.

[1] Fimmel E., Michel Ch. J., Pirot, F., Sereni, J.-S., and Strüngmann L.: Circular Codes over Finite Alphabets, work in progress.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The evolutionary influence of the aminoacyl-tRNA-synthetase feedback loop

Hanna Schumacher
University of Applied Sciences Mannheim, Faculty of Computer Science
Paul-Wittsack-Straße 10, 68163 Mannheim, Germany,
hanna@schumacher-home.de

Feedback loops in biology exist on many different levels and help to maintain homeostasis. A common feedback loop example are predator and prey populations, but this presentation is about a feedback loop on a much lower level, on the level of protein biosynthesis. An essential part of the translation process are aminoacyl-tRNA-synthetases (aaRSs). They combine tRNAs with amino acids and with this they define the genetic code itself. On one hand aaRSs have a big influence on the translation products, on the other hand they are translation products themselves. Without considering evolutionary mechanisms like natural selection, genetic drift and mutations a model for the aaRS feedback loop was created. This allows to analyze the influence of the feedback loop undisturbed by some of the main drivers of evolution. The first part of the presentation is the explanation of this model by using a simplified example. Based on this model a computer simulation was programmed. The code for the simulation and the code for analyzing the results with RStudio can be found on Github*. A short explanation is given as the second part of the presentation. The simulation has been run with various parameter combinations and the resulting code table fitness values were analyzed. The formula for the fitness calculation considers codon unambiguousness and number of translated amino acids. The results of the analysis are presented as last part of the presentation. The analysis is based on statistical tests and the usage of the J48 data mining algorithm of the Weka-System. All in all, a new approach on the field of genetic code evolution is presented together with the tools for further exploration and first results.

* https://github.com/Thalla/GeneticSimulations and https://github.com/Thalla/GeneticSimulationsR

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Spontaneous evolution of circular codes in theoretical minimal RNA rings

Jacques Demongeot[1], Hervé Seligmann[2]
[1]Université Grenoble Alpes, Faculty of Medicine, Laboratory AGEIS EA 7407,
Team Tools for e-Gnosis Medical, F-38700 La Tronche, France.
[2]The National Natural History Collections,
The Hebrew University of Jerusalem, IL-91904 Jerusalem, Israel.
podarcissicula@gmail.com

Small circular RNAs occur in many cells. Theoretical considerations designed to mimick primordial RNAs designed 25 theoretical, minimal 22-nucleotide long circular RNAs (RNA rings) forming stem-loop hairpins, including one codon per amino acid, a start and a stop codon. These RNA rings resemble consensus tRNAs, whose predicted anticodons assign to each RNA ring a potential cognate amino acid. Assuming dual translator and messenger roles, three consecutive translation rounds produce 21-residue-long peptides, seven codons at each round. In these conditions, steric hindrances between tRNA(-like) translators competing for partially overlapping nucleotide triplets can be avoided if none of the seven codons produces by circular permutation (position 1->3, or position 3->1) any of the six others. This non-permutability of codon sets defines potential circular codes regulating translational frame. A near-universal maximal self-complementary circular code exists in reading frames of natural genes, and permutations of this circular code define maximal circular codes in each +1, +2 gene frames. Scaling RNA rings according to the genetic code inclusion order of their cognate amino acid, codon numbers belonging to the natural frame +1 circular code X1 decrease with cognate amino acid inclusion order, those belonging to the natural frame 0 circular code X0 increase. RNA rings with early cognates apparently reflect pre-(tRNA-like)-adaptor translation by direct codon-amino acid affinity with partially overlapping consecutive codons where X1 regulated translation. Translation of non-overlapping consecutive codons evolved in parallel with RNA rings' cognate inclusion in the genetic code and with X0. Hence the complex "multi-frame" natural circular codes potentially evolved spontaneously from small coding circular RNAs mimicked by theoretical minimal RNA rings. Modern reading frames evolved from earlier reading frames corresponding to modern +1 frames.
Key words: tRNA synthetase class; cloverleaf secondary structure; anticodon loop; comma-free codes.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

A helix, a spiral, a torus: a code-behind beautiful higher-order structures of DNA

Kirti Prakash
Department of Paediatrics, Wellcome-MRC Cambridge Stem Cell Institute
University of Cambridge, Cambridge, CB20QQ, UK
kp511@cam.ac.uk, kirtiprakash2.71@gmail.com

There is sheer beauty to the helical structure of DNA (1, 2, 3). A helix is probably the most refined structure by which two polymers may couple and provide a pairing mechanism for the most efficient form of replication. The helix wraps around octamers of histone proteins to form nucleosomes in a fascinating beads-on-a-string structure (4). DNA, nucleosomes, nucleus, cells, tissues, organs, organisms, species are not random assignments of evolution without design but factual realities arising from millions of years of tinkering (5, 6). Hierarchy, whether natural or social, helps to achieve an appropriate classification of the components of various systems based on space, time, topology or power. By recursively distributing energy to the smaller building blocks, hierarchy maintains a necessary genetic structure for information flow.
Here, I describe new putative building blocks of DNA that are functionally distinct with unique topological features (7, 8). I put forward a theory based on 'histone code' on how these building blocks can generate a model by which DNA folds hierarchically (9, 10, 11).

1. Franklin, Rosalind E., and Raymond G. Gosling. "Molecular configuration in sodium thymonucleate." Nature 171 (1953): 740-741.
2. Watson, James D., and Francis HC Crick. "Molecular structure of nucleic acids." Nature 171.4356 (1953): 737-738.
3. Wilkins, Maurice Hugh Frederick, Alex R. Stokes, and Herbert R. Wilson. "Molecular structure of nucleic acids: molecular structure of deoxypentose nucleic acids." Nature 171.4356 (1953): 738.
4. Kornberg, Roger D., and Yahli Lorch. "Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome." Cell 98.3 (1999): 285-294.
5. Barbieri, Marcello, ed. The codes of life: the rules of macroevolution. Vol. 1. Springer Science & Business Media, 2007.
6. Barbieri, Marcello. "From biosemiotics to code biology." Biological Theory 9.2 (2014): 239-249.
7. Prakash, Kirti, et al. "Superresolution imaging reveals structurally distinct periodic patterns of chromatin along pachytene chromosomes." Proceedings of the National Academy of Sciences 112.47 (2015): 14635-14640.
8. Kirmes, I., Szczurek, A., Prakash, K., Charapitsa, I., Heiser, C., Musheev, M., ... & Cremer, C. (2015). A transient ischemic environment induces reversible compaction of chromatin. Genome biology, 16(1), 246.
9. Jenuwein, Thomas, and C. David Allis. "Translating the histone code." Science 293.5532 (2001): 1074-1080.
10. Prakash, Kirti, and David Fournier. "Evidence for the implication of the histone code in building the genome structure." Biosystems (2017).
11. Prakash, Kirti, and David Fournier. "Histone code and higher-order chromatin folding: A hypothesis." Genomics and Computational Biology 3.2 (2017): 41.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Assimilation of NonReduction Theorem in Context of Topological Explanations

Lukáš Zámečník, Ľudmila Lacková
Department of General Linguistics, Palacký University
Olomouc, Czech Republic
ludmila.lac@gmail.com

The aim of this presentation is to propose potential varieties of non-mechanistic explanations in the context of biology. We try to demonstrate the validity of Peirce´s iconic logic in the context of the current topological explanations in the philosophy of science. Arguing that the logical system of Existential Graphs is not dissimilar from the contemporary topological approaches, we recognise Peirce´s iconic as a valid method of scientific representations. We base our argumentation on the connection between the iconic logic and the so-called NonReduction Theorem and we illustrate our assumptions with aid of examples taken from biology. Concretely we deomonstrate the NonReduction Theorem with aid of examples of protein folding and genotype → phenotype maps. The presentation has the following scopes. Firstly, we introduce the notion of NonReduction Theorem. We show how to understand triadic predicates by means of the Peircean iconic (diagrammatic) logical system. Secondly we extend the notion of NonReduction Theorem to a more general philosophical context, namely to the context of liberal naturalism within the philosophy of science. Thirdly, in connection to the previous sections, dyadic and triadic conceptions of logical relations are compared in the particular context of biology. The notion of topological explanation is described. The aim of this paper is to bring Peircean iconic logic (Existential Graphs) to the contemporary debates on topological explanations in the philosophy of science, proving in this way the validity of Peirce's iconic logic in contemporary science.

Keywords
Beta Graphs, iconic logic, protein folding, topological explanations

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The standard genetic code is not as robust to amino acid replacements as its many alternative variants

Małgorzata Wnętrzak, Paweł Błażej, Dorota Mackiewicz,
Przemysław Gagat, Paweł Mackiewicz
Department of Bioinformatics and Genomics, Faculty of Biotechnology,
University of Wrocław, Poland
malgorzata.wnetrzak@uwr.edu.pl

We checked the robustness of the standard genetic code (SGC) and its known alternative variants, in terms of the consequences of amino acid replacements. Moreover, we compared the results with the properties of all the theoretical genetic codes that differ from the SGC by one, two, or three codon assignments.
Even though the SGC is closer to the best theoretical codes than to the worst ones, considering the robustness to changes of amino acids in proteins, there are still many possible codes that minimize the effects of such changes to a greater extent than the SGC. It is also interesting that many types of codon reassignments observed in the existing alternative genetic codes contribute significantly to the increase in the robustness to amino acid replacements, in comparison with the SGC. According to our optimization criteria, 18 out of 21 considered alternative codes perform better than the SGC.
These results indicate that not all codon reassignments in the alternative codes are neutral and some of them could have been selected to reduce harmful effects of mutations on the protein-coding sequences and translational errors. These findings also show that only a few codon reassignments are necessary to substantially improve the robustness of the SGC.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The time of the living – on the poetics of recognition

Marcella Faria, PhD
Dactyl Foundation for the Arts and Humanities
64 Grand Street, New York, NY 10013 USA
almeidapradomarcella@gmail.com

Living systems – either real or fictional – display a precise sense of time. In processes as diverse as, the evolution of species; the development of embryos; the growth of cells; and the journey of literary heroes, time acts as a differentiation agent. It generates specifity, discriminatory competencies, patterns. At the same time, living things, as historic as they might be, will always be immersed in the present of a mutable landscape. Boundary conditions, *i.e.* the constrains imposed to a system on its edges, are always changing. Living beings must continuously adapt to contingent resources inside their history, crafting narratives. In our view, this ability to build narratives, as a distinctive feature reuniting biology and literature, emerges by a poetic use of material resources. What we shall illustrate by a brief discussion of two theoretical frameworks:

1) The Sciences of Recognition, as postulated by the neuroscientist Gerald Edelman in 2004;
2) The Theory of Poetic Action, as proposed by Paul Valéry in the late 1930s'

Edelman claims that in evolution, embryology, immunology and neurobiology, a precise setting of biological rules emerges by selection acting over time on variable populations of molecules, cells, and organisms. These two notions, i.e. variation and selection, are at work at all levels in each scenario. We will translate Edelman's principles of recognition in terms of Valéry's theory of poetic action.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany), 3-7 June 2019

Evolution of the Genetic Code: the Ambiguity-reduction Theory

Marcello Barbieri
Dipartimento di Morfologia ed Embriologia
Via Fossato di Mortara 64a, 44121 Ferrara, Italy
brr@unife.it

The experimental evidence has shown that the genetic code is based on *arbitrary*, or *conventional*, rules, in the sense that any codon can be associated to any amino acid, and this means that there is no deterministic link between them. This is in sharp contrast with the two most popular theories on the genetic code – the *stereochemical theory* and the *coevolution theory* – because both of them assume that the rules of the genetic code were determined by chemistry, either by stereochemical affinities or by metabolic reactions. The discovery that the genetic code is based on arbitrary rules, on the other hand, raised a formidable problem: how can such rules exist in Nature? In order to deal with this problem, it has been pointed out that the rules of an arbitrary code could not come fully formed into existence. The first genetic code, in other words, was necessarily ambiguous, and its evolution took place with a mechanism that systematically reduced its ambiguity and eventually removed it altogether. The concept of ambiguity-reduction has been repeatedly mentioned in the scientific literature, but for a long time it has remained only an abstract possibility because no model was proposed for its mechanism. The first paper that described that mechanism was published with the name of *ribosome-oriented model* in order to underline the key role that the ribosomal proteins had in that process, but later on it became clear that other factors had to be to be taken into account. This is why the *ribosome-oriented model* had to be extended and here a more general version is proposed with the name of *ambiguity-reduction theory*.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Composition principles of long DNA sequences compared with random sequences

Markus Gumbel[1]
Competence Center for Algorithmic and Mathematical Methods
in Biology, Biotechnology and Medicine, Mannheim University of Applied Sciences
m.gumbel@hs-mannheim.de
joined work with
Sergey Petoukhov[2] and Elena Fimmel[1]

The discovery of basic mechanisms of the order of bases in long DNA sequences – like in a complete chromosome – is one of the crucial tasks in biological systems. In this presentation, the frequencies of the base composition found in biological and Markov-like randomly generated similar sequences are analyzed and compared. The frequencies were calculated for the entire sequence as well as in subsequences of these, whereby a subsequence contains only those nucleotides that have a fixed distance of *n* bases. This approach is to some extend a generalization of Chargaff's Second Parity Rule [1, 2]. Our observations with a distance of up to *n* = 50 bases have proven that Chargaff's Second Parity Rule also applies to large distances. Also, no significant differences between the real and simulated sequences could be observed. Next, the frequency analysis was applied to partitions of the entire chromosomal sequence. Here the analysis revealed that the frequencies of the base composition vary strongly in the chromosome's regions – far more than it would be expected due to regions with a high gene density or CpG islands. And most strikingly, this could not be observed in the randomly generated sequences. Eventually, this talk will discuss when a DNA sequence can be considered as "long" and how to quantify the choice of length.

[1] Chargaff E, Lipshitz R, Green C: (1952) Composition of the deoxypentose nucleic acids of four genera of sea-urchin. J Biol Chem. 195 (1): 155–160.
[2] Petoukhov S.: (2018) The rules of long DNA-sequences and tetra-groups of oligonucleotides, arXiv:1709.04943v5[q-bio.OT].

[1] Competence Center for Algorithmic and Mathematical Methods in Biology, Biotechnology and Medicine, Mannheim University of Applied Sciences (m.gumbel@hs-mannheim.de, e.fimmel@hs-mannheim.de)
[2]Laboratory of biomechanical systems, Russian Academy of Sciences, Moscow, Russia (spetoukhov@gmail.com).

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Genetic code engineering to realize genetic firewalls: a bioinformatics evaluation of potential development pathways

Markus Schmidt,
Biofaction KG, Vienna, Austria
schmidt@biofaction.com

The genetic code mediates the translation from the nucleic acid alphabet (consisting of 4 bases ATGC in DNA, or AUGC in RNA) to the amino acid alphabet (consisting of 20 proteinogenic amino acids) via triplets (three bases). While most life forms on Earth use the standard genetic code, 30 alternative codes have been identified so far in Nature, differing only in a few triplets from the standard code. Genetic Code Engineering (GCE) is a branch of synthetic and xenobiology that aims to design and implement genetic codes not found in nature. One of the reasons to do GCE is the expected shift in the cellular information processing capabilities leading to genetic firewalls reducing or eliminating altogether the possibility of horizontal gene transfer. This type of semantic containment might become a novel biosafety device, separating natural from synthetic life forms. Based on a recently developed metric to measure the distance between any two genetic codes (Schmidt 2018), viable development pathways towards independent information processing islands are analysed with bioinformatics methods. The GCE pathways analysed here all involve gradual modification of triplet assignments, either by codon blocks or individual codons. In contrast to previous assessments (for an exception see: Błazej et al. 2018) the numerical evaluation presented here also attempts to cover the reassignment of STOP codons (in particular by assigning artificial code damage values), a process frequently observed in alternative natural codes. Depending on the algorithm (blocks versus single triplets) and specifications (e.g. for STOP reassignments or internal mutational robustness), different anti-codes (that are codes that are most dissimilar to the e.g. standard code) are presented. The different approaches and results are discussed regarding their efficiency to create genetic firewalls and their potential to be realized in actual lab experiments.

Keywords: Synthetic biology, xenobiology, biosafety, biocontainment,

Schmidt 2018. Introducing a metric to generate alternative genetic codes for the implementation of (a) genetic firewall(s). Fifth International Conference in Code Biology. Granada, Spain. 5-9 June 2018

Błazej, P., Wnetrzak, M. and Mackiewicz, P. 2018. The Importance of Changes Observed in the Alternative Genetic Codes. In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018) - Volume 3: BIOINFORMATICS, pages 154-159

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Possible advantages of k-circular codes in the mitochondrial genetic code

Martin Starman
Mannheim University of Applied Sciences,
Paul-Wittsack-Str. 10, 68163 Mannheim, Germany
&
Université de Strasbourg,
CNRS 300 Boulevard Sébastien Brant, 67400 Illkirch, France
martin.starman@etu.unistra.fr

The first idea of an error detecting mechanism in the mitochondrial genetic code was proposed by Crick et al. in 1957 [3]. This proposal was based on the so-called comma-free codes. Further research shortly after the publication [3] has shown that the error detection in the mitochondrial genetic code cannot be explained so easily. Yet, the X code discovered by Michel and Arquès through an intensive statistical investigation has given new impetus to this topic [1]. This code has all the properties of a circular code, which is from the same family as comma-free codes, but slightly less restrictive [2, p. 233]. The code is a set of the 20 most occurring codons in bacteria, archaea, eukaryotes, plasmids and viruses [1].
Even though this discovery has been a breakthrough, not all problems have been solved. This fact can be explained by the incomplete knowledge of the circular code family. Though the comma-free codes had a lot of attention, the rest of the circular code family was left behind [4][5]. This talk introduces the so-called k-circular codes. These codes are even less restrictive than circular codes. Depending on the actual sequence, these codes do not necessarily discover an error. While they are less reliable, they do allow a greater maximal length then the standard circular codes. This advantage of k-circular code could be a major lead to explain the snippets between the sections with codons from the X code in real live sequences.

References

[1] D. G. Arquès and C. J. Michel. A complementary circular code in the protein coding genes. *J. Theor. Biol*., 182(1): 45–58, 1996.
[2] J. Berstel, D. Perrin, and C. Reutenauer. *Codes and automata*. 2005.
[3] F. H. Crick, J. S. Griffith, and L. E. Orgel. Codes without commas. *Proc. Natl. Acad. Sci. U.S.A.,* 43(5):416–421, 2003.
[4] S. W. Golomb, L. R. Welch, and M. Delbrück. *Construction and properties of comma-free codes.* i kommission hos Ejnar Munksgaard, 1958.
[5] N. H. Lam. Completing comma-free codes. *Theoretical computer science,* 301(1):399–415, 2003.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The search image as link between sensation, perception and action

Morten Tønnessen
University of Stavanger
Nådlandsbråtet 25, 4034 Stavanger, Norway
mortentoennessen@gmail.com

This presentation is based on a published article, Tønnessen 2018.

In this paper, it is argued that Jakob von Uexküll's "search image" notion, the original version of this notion within ethology, is still of use. A search image, in Uexküll's sense, is an imagined object that an organism has in mind when it searches for something. Uexküll's conception of the search image is useful both for understanding the theoretical context of contemporary notions of search images, and with an eye to envisioning possible future developments of the idea. Uexküll's classical notion differs from contemporary versions in that it has a wider application, and is therefore of greater relevance to theoretical biology and cognitive science. It constituted an integral part of his ground-breaking Umwelt theory, stressing the fundamental plasticity of the Umwelt, the subjective lifeworld of an animal or human subject.
In a contemporary development of Umwelt theory, expressed by the tripartite Umwelt model, the search image notion represents a key connection between the directly experienced core Umwelt and the mediated Umwelt. However, the key function that schemata have in cognitive processes is also the starting point for mistakes in perception.
This article details both the constructive function search images have in animal and human perception, and the mismatches in perception they can lead to. It also explains how the existence of search images can help explain puzzles concerning subjective and neural time in contemporary cognitive science.

Reference
Tønnessen, Morten 2018. The search image as a link between sensation, perception and action. Contribution to the first special issue in Code Biology, *BioSystems* 164 (2018): 138–146. Published online November 4th 2017. DOI: 10.1016/j.biosystems.2017.10.016.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

In search of a primitive signaling code

Nadir M. Maraldi
Department of Biomedical and Neuromotor Sciences, University of Bologna, Italy
nmaraldi@area.bo.cnr.it

Cells must have preceded by simpler chemical systems (protocells) that had the capacity of a spontaneous self-assembly process and the ability to confine chemical reaction networks together with a form of information. The presence of lipid molecules in the early Earth conditions is sufficient to ensure the occurrence of spontaneous self-assembly processes, not defined by genetic information, but related to their chemical amphiphilic nature. Ribozymes are plausible molecules for early life, being the first small polynucleotides made up of random oligomers or formed by non-enzymatic template copying. Compartmentalization represents a strategy for the evolution of ribozymes; the attachment of ribozymes to surfaces, such as formed by lipid micellar aggregates may be particular relevant if the surface itself catalyzes RNA polymerization. It is conceivable that the transition from pre-biotic molecular aggregates to cellular life required the coevolution of the *RNA world*, capable of synthesizing specific, instead of statistical proteins, and of the *Lipid world*, with a transition from micellar aggregates to semipermeable vesicles. Small molecules available in the prebiotic inventory might promote RNA stability and the evolution of hydrophobic micellar aggregates into membrane-delimited vesicles. The transition from ribozymes catalyzing the assembly of statistical polypeptides to the synthesis of proteins, required the appearance of the *genetic code*; the transition from hydrophobic platforms favoring the stability of ribozymes and of nascent polypeptides to the selective transport of reagents through a membrane, required the appearance of the *signal transduction code.*

Here I present evidences on the presence of traces of the evolution of a signal transduction system in extant cells, which utilize a phosphoinositide signaling system located both at nucleoplasmic level as well as at the plasma membrane, based on the very same molecules but responding to different rules. The signals from the environment, however, were unable to bypass the hydrophobic self-assembled aggregates, owing to the absence of a system of active transport based on 3D complex protein machinery, but can affect the arrangement of these aggregates which, in turn, can act on the ribozymes through a rudimental signaling system, whose traces are still available in the signaling system based on phosphoinositides acting at the nuclear level (Maraldi, 2008). The evolutionary significance of the organic codes, besides the genetic code, has been widely recognized (Barbieri, 2003); in particular the effects that external signals could have on a cell do not depend on the information that they carry, but on the meaning that cells give them owing a set of rules (signal transduction code) (Barbieri, 2016).

Lipids, polypeptides, but not proteins acting as enzymes, and polynucleotides have been conceivably formed in pre-biotic conditions. In order to evolve towards a protocellular state, autoreplicating ribozymes, capable to catalyze polypeptide synthesis, might be temporarily isolated with respect to the environment. Compartmentalization, being a prerequisite of cellular life, might be acquired by a coevolving process, due to autoassembly mechanisms of lipids, induced by hydrophobic forces.

Autopoietic computation and Shcherbak numbers

Nataša Ž. Mišić
Research and Development Institute Lola Ltd, Belgrade, Serbia
nmisic@rcub.bg.ac.rs

The various conceptual frameworks for the organization of living systems have been mainly developed independently, which implied many similarities and differences between them. However, the central idea common to all these theories is the operational *closure*, particularly rigorously developed in the framework of *(M,R) systems* [1] and *Autopoiesis* [2], where Autopoiesis can be regarded as a subset of (M,R) systems [3]. Autopoietic systems possess not only a significant place in theoretical biology but also in an abstract formalization of natural biocomputing and its technical implementation in the form of nature-inspired (bio)computing ("soft" computing) [4]. Here, some analogical properties between a computation within Autopoietic systems and within a special class of numbers resulted from the arithmetical regularities inside the Standard Genetic Code (SGC), so-called the *Shcherbak numbers* [5,6], are highlighted.

Autopoietic (self-producing) system, and thus Autopoietic computation, is distinguished by at least five specific features and characteristics: 1) maintains its circular organization invariant (an operational closure by circular causality or recursivity); 2) depends merely upon the relations between its components, not directly on the nature of the components themselves (a relational property); 3) couples a purely relational property with a topological property (an autonomous unity); 4) its boundaries are continuously self-produced as a result of self-determined internal dynamics (a self-referential property), and 5) its future state is determined not solely by a present state and the perturbation but also by a future state (an anticipatory property). A deeper analysis of some mathematical properties of Shcherbak numbers (their relation to the cyclic structures and regular multiplying [5,6] as well as to the self-referential generating formulas, self-similar nested structures, decimal scaling, recurrent formulas, and figurative numbers [6]) reveals their compatibility with some concepts of Autopoietic computation and therefore provides insights that can contribute to a better understanding of the mentioned features and characteristics of Autopoietic systems and the arithmetical nature of SGC.

The authors acknowledge support from the Ministry of Education, Science and Technological Development of Serbia (Projects: TR 32040 and TR 35023).

1. R. Rosen, *Bull. Math. Biophys* **20** (1958) 245–260; **21**, 109–128; *Anticipatory Systems*, Pergamon, 1985; *Life Itself*, Columbia University Press, 1991.
2. F. Varela, H.R. Maturana, R. Uribe, *Biosystems* **5** (1974) 187–196; F. Varela, *Principles of biological autonomy*, Elsevier, 1979; H. Maturana, F. Varela, *Autopoiesis and cognition: The realization of the living*. Springer, 1980.
3. J.C. Letelier, G. Marín, J. Mpodozis, *J. Theor. Biol.* **222** (2003) 261–272; J.C. Letelier, M.L. Cárdenas, A. Cornish-Bowden, *J. Theor. Biol.* **286** (2011) 100–113.
4. J.C. Letelier, G. Marín, J. Mpodozis, In *Soft Computing and Industry*, R. Roy et al. (Eds). Springer, 2002, 67–80; J.C. Letelier et al., In *Algebraic biology*, H. Anai, K. Horimoto (Eds.), Universal Academy Press, 2005, 115–126.
5. V.I. Shcherbak, *J. Theor. Biol.* **166** (1994) 475–477; *Biosystems* **70** (2003) 187–209; In: *The Codes of Life*, M. Barbieri (Ed.), Springer, 2008, 153–188.
6. N.Ž. Mišić, *Proc. NEUREL*, 2010, 97–102; *Neuroquantology* **9** (2011) 702–715; *Proc. TABIS*, 2014, 101–123; *FU Phys. Chem. Tech.* **14** (2016) 275–298.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Genetic Code Engineering, Artificial Organisms and Biocontainment

Ned Budisa
[a]Berlin Institute of Technology/TU Berlin, Department of Chemistry, Biocatalysis Group
Müller-Breslau-Straße 10, D-10623 Berlin, Germany,
budisa@chem.tu-berlin.de
[a]University of Manitoba, Department of Chemistry & Microbiology,
Chair of Chemical Synthetic Biology 144 Dysart Rd, R3T 2N2 Winnipeg, MB, Canada

Life on earth is a unity owing to the existence of the universal genetic code, i.e. as the genetic code for all organisms is basically the same - all living things use the same "genetic language". This biological framework allows the translation of the genetic message written in DNA into life-sustaining proteins. Therefore, the genetic code can be referred to as the *lingua franca* of life on earth, which allows for the maintenance of universal biochemistry. The standard translational apparatus also requires a uniform genetic code structure. These are prerequisites for the dissemination of biological novelties in biosphere by horizontal gene transfer between all life forms. Due to the code universality, the genetic information is always translated into the same protein in the same way, regardless of the host organism. Biocontainment arises when a genetically or chemically altered organism can operate under a different genetic code, either by codon reassignment or by changing the decoding rules. For example, synthetic biology opens up the possibility of minimizing the risk of virus or bacteriophage infection during culture, as synthetic cells (with altered genetic codes) are no longer suitable as hosts for viruses and phages. This means that artificial cells have a higher resistance through so-called "semantic containment". This makes a completely new biological world conceivable and plausible. The design of genetically modified organisms (in the context of classical genetics) is only the beginning of a long road in search of reliable methods for the design and evolution of artificial biodiversity while preserving the “old” natural world. Therefore, an important task for engineering biology is to use directed evolution in order to create viable and robust synthetic cells with alternative biological coding that can grow and replicate indefinitely in isolation from natural species.

References

[1] Kubyshkin, V., Acevedo-Rocha, C. and Budisa N. On universal coding events in protein biogenesis. *Biosystems*. 2017, *S*0303-2647_30278-2.doi: 10.1016/j.biosystems.2017.10.004.
[2] Kubyshkin, V. and Budisa N. Synthetic alienation of microbial organisms by using genetic code engineering: why and how? *Biotechnol. J*. 2017, *12* (8), doi: 10.1002/biot.201600097.
[3] Aggarwal, N., Bandhu, A.V. and Sengupta, S. Finite population analysis of the effect of horizontal gene transfer on the origin of a universal and optimal genetic code. *Phys Biol.* 2016, 13(3):036007. doi: 10.1088/1478-3975/13/3/036007.
[4] Acevedo-Rocha, C. and Budisa, N. Xenomicrobiology: a roadmap for genetic code engineering. *Microb. Biotechnol*. 2016, *9*, 666-676, doi: 10.1111/1751-7915.12398.
[5] Budisa, N. Xenobiology, New-to-Nature Synthetic Cells and Genetic Firewall. *Curr Org Chem* 2014, *18*, 936–943, doi: 10.2174/1385272818081406161543O1.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The complementarity code: The tool for protein structure, function and evolution analysis

Nikola Štambuk[1] and Paško Konjevoda[1]
[1]Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia
stambuk@irb.hr pkonjev@irb.hr

The concept of complementary peptide interaction was developed in the early 1980s by Robert Scott Root-Bernstein and James Edwin Blalock [1]. Sense and antisense peptides specified by the complementary DNA and RNA sequences have been investigated by a number of research groups and authors, and during the last four decades this approach was experimentally verified for more than 50 ligand–acceptor (receptor) systems [1]. Recent investigations have shown that the interaction of sense and antisense peptides can be described with the set of rules that combines the following elements: of amino acid physico-chemical properties, bidirectional coding, and stereochemical interaction [1]. This natural genetic coding algorithm has been named *the complementarity code*. It is a sequence code that uses the elements of the Standard Genetic Code (SGC) and recently-described Carter-Wolfenden tRNA acceptor-stem code (tRNA ASC) [1,2], and possesses all the criteria for an organic code [3]. The characteristics of *the complementarity code* are:

1. It translates more antisense amino acid pairs in the 3'→5' direction (27) then in the 5'→3' direction (52) and sets up *the complementarity code* table [1];
2. interactions of complementary amino acid pairs, in terms of hydrophobicity and lipophilicity, strongly depend on the central purine bases of the mRNA codons and its pyrimidine complements of the tRNA anticodons—in both translation directions; and
3. translation process specifies 3 clusters of sense–antisense amino acids that pair according to the character of the second SGC base columns: polar (2^{nd} A) with nonpolar (2^{nd} U), and neutral (2^{nd} G) with neutral (2^{nd} C)—irrespective of physicochemical parameters (hydrophobicity or lipophilicty), and the direction of reading (3'→5' or 5'→3').

The results show that the values of amino acid hydrophobic moments follow *the complementarty code* rules, with respect to the protein secondary structure determination [1,4,5]. Additionally, Carter-Wolfenden hydrophobicity and lipophilicty parameters enable the calculation of distances between amino acid pairs, which could be used as a supplement to the PAM and BLOSUM Substitution Matrices for scoring alignments between evolutionarily divergent protein sequences [1,6]. *The complementarity code* could have both theoretical and practical applications in the fields of protein structural and functional analysis. Other types of application include: the investigation of protein and nucleotide sequences evolution, virtual screening of peptide ligands, and artificial (*de novo*) protein construction.

References

1. Štambuk, N., Konjevoda, P., Turčić, P., et al., 2018. BioSystems 164, 199–216
2. Carter, C.W., Wolfenden, R., 2015. Proc. Natl. Acad. Sci. USA. 112, 7489–7494.
3. Barbieri, M., 2015. Code Biology—A New Science of Life, Springer, Cham.
4. Štambuk, N., Konjevoda, P., 2017. Information 8, E1–E29; doi:10.3390/info8010029.
5. Štambuk, N., Konjevoda, P., Manojlović, Z., et al., 2016. Symmetry: Cult. Sci. 27, 163–171.
6. Štambuk, N., Konjevoda, P., Manojlović, Z., 2016. Lecture Notes in Computer Science, vol 9656. Springer, Cham, pp. 358–365.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Lexicon analysis of the sign set in the genomic codes of noncoding regulatory regions

O. Paredes, R. Romo-Vázquez, J. A. Morales
Departamento de Ciencias Computacionales, CUCEI,
Universidad de Guadalajara, Mexico
omar.paredes@alumnos.udg.mx

Code Biology is a field that formulates that life is not merely chemistry and information but organic codes. An organic code is a set of rules that maps the elements of a set (sign set) into another (significance set), where both sets are independent between them and the rules are specific but arbitrary at the same time [1]. A widely study organic codes is the genetic code, that is an example of a broader set of organic codes called genomic codes. A genomic code is an organic code that maps patterns of nucleotides into a significance set, in the case of the genetic codes are the set of the 20 aminoacids. Hofmeyr [2] states that the methodology to determine if an organic code consists in i. demonstrate that the code links two independent sets, from biomolecules (sign set), in this case to, biological effects (significance set); ii. identify the adaptor that recognizes and decodes the organic code; and iii. prove the arbitrariness nature of the rules in the organic code. In this work, we analyze the genomic lexicon to establish the sign set of nucleotide patterns that maps into the significance set of the regulatory regions proposed by the Epigenomic Roadmap project and the links between both sets in order to evaluate the regulatory codes as an organic code. We implement techniques of genomic signal processing and natural language to determine the size of the genomic codes, evaluating the symbolic, numeric and sparse representations of the noncoding regulatory sequences and identifying the lexicon that characterizes the regulatory regions. These results suggest that the regulatory regions are not only determined by the histone code but to another layer of information coded in the genomic sequence and hint to a multicode phenomena in complex organisms as Eukarya organism.

1- Barbieri M: What is Code Biology? *Biosystems* 2018, 164:1–10.
2- Hofmeyr J-HS: The first Special Issue on code biology – A bird's-eye view. *Biosystems* 2018, 164:11–15.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Genetic Code, Symmetry and Minimum Principle

Paul Sorba
Laboratoire d'Annecy-le-Vieux de Physique Théorique
CNRS and Université de Savoie
paul.sorba@lapth.cnrs.fr

A model for the Genetic Code based on continuous symmetries and entitled the ''*Crystal Basis Model*" has been proposed a few years ago. We will present a review of the model, of some of its first applications as well as of its recent developments. More precisely, after a motivated presentation of our mathematical model, we illustrate its pertinence by applying it for the elaboration and veryfication of sum rules for codon usage probabilities, as well as for establishing relations and some predictions between physical-chemical properties of amino-acids. Then, defining in this context a ''*bio-spin*" structure for the nucleotides and codons, the interaction between a couple of codon-anticodon can simply be represented by a (bio) spin-spin potential. This approach will constitute the second part of the talk where, imposing the minimum energy principle, an analysis of the evolution of the genetic code can be performed with good agreement with the generally accepted scheme. A more precise study of this interaction model provides informations on codon bias, consistent with data.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Theoretical model of genetic code structure evolution

Paweł Błażej, Małgorzata Wnętrzak, Paweł Mackiewicz
Department of Genomics, Faculty of Biotechnology,
University of Wrocław, Poland
blazej.pawel@gmail.com

In this work, we present a theoretical model of genetic code evolution where potential genetic codes evolved from system of ambiguous codon assignments to the set of rules with reduced level of uncertainty. Especially, we are interested in testing some factors which would be responsible for compact structure of the standard genetic code. In this context, we consider three types of codon reading systems creating codon blocks assigned to respective amino acid. To do so, we run simulations where misstranslations can act only on fixed positions in codon.
This process starts from randomly generated genetic codes with ambiguous codon assignments and they are selected to improve coding ambiguity and their robustness against misreading. The results indicate that the evolution of genetic code carried out under restrictions similar to those which are observed in the standard genetic code is responsible for finding solutions with the best quality.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The Analogy between Cellular Codes and Computer Codes

Pedro Marijuán
Bioinformation Group, Aragon Institute of Health Science, Zaragoza, Spain
pcmarijuan.iacs@aragon.es

The analogy between Cellular Codes and Computer Codes has been often considered from highly theoretical points of view, such as Turing Machine, Shannon's theory, von Neumann architecture, semiosis, etc. Conversely, the view developed here will be closer to reverse engineering, i.e., attempting a bottom-up interpretation of the mechanisms and dynamics that we can directly appreciate in these processing systems. Both cells and computers will appear as full of codes. In the case of computers, there is a universal reliance on the binary nature of all the circulating instructions, data, and signals. All the massive exchanges of information circulate as binary strings within a complex architecture or structure of interrelationships among the main components. Each string has a specific structure of codes impresses upon it. Given that the main components have different architectures and processing properties, they need specific codes to enter their inputs and send their outputs via the common bus system that interconnects them. Further, a full hierarchy of codes is established in between the different processing levels: microcodes, machine codes, peripheric codes, instruction codes, language codes, user codes, interconnection codes, networking codes…
In the case of cells, there is a water matrix that provides "wiredless" massive interconnection to all other components via Brownian motion. Molecular recognition becomes the key functional phenomenon from which the whole cellular dynamics is built. There is specificity of interactions: exact recognition in myriads of different kinds of contacts between molecular partners. Combinatoric interactions are organized by means of "lock & key" combinations of covalent bonds, ionic bonds, hydrogen bonds, hydrophobic / hydrophilic forces, dipole forces, van der Waals forces, etc. There emerge three main classes of molecular recognition that correspond to main informational architectures of the cell: the sequential (DNA, RNA), the processual (enzymes/proteins), the structural (membranes).
Symmetry considerations: they respectively mean recognitions of identity (structural), complementarity (sequential), supplementarity (processual). Thereafter, the self-construction of the cellular system via the central genetic code actually implies a number of *microcodes*, whenever and wherever information has to travel with specificity from a particular class of architecture to another—i.e., when sequential motifs have to be copied or be recognized by enzyme processors, when proteins have to recognize each other, or when organelle structures have to be built or modified. Whatever functional operation has to be performed by the synthesized enzymes and proteins, it implies the generation of new recognition codes among the heterogeneous architectural partners involved (e.g., for protein modification, transportation, splicing, complexes, etc.), including the final degradation in the proteasome as well. All complex functions would partake of similar basic designs based on an assemblage of low-level and middle-level codes for the matching between the different architectures involved.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Molecular Codes and Self-maintaining Sets

Peter Dittrich
Bio Systems Analysis Group, Institute of Computer Science
Friedrich-Schiller-Universität, Jena, Germany
peter.dittrich@uni-jena.de

Previously, a formal method to assess the semantic capacity of a chemical reaction network to process "meaningful" information has been suggested (Görlich/Dittrich, PLoS ONE, 8(1), e54694, 2013). The basic idea is to measure how easy it is to implement with this network a molecular code, which is an arbitrary (contingent) mapping between species, that is, a mapping that cannot be inferred from knowing the species alone. A preliminary computational analysis of various chemical systems revealed a quite large spectrum of different semantic capacities. The hypothesis has been derived that life over the course of evolution is gaining access to chemistries with increasing semantic capacity. A recent study revealed molecular codes in metabolic network models; however basically each code contained an ubiquitous species like water, so that the contingency of the codes found is questionable. As a consequence, the theory has to be refined along the following two questions: (1) Which component of a (molecular) system should be regarded as a sign, meaning, or code-maker, respectively? This would be important for deriving experimental procedures and algorithmic tools for finding and characterizing molecular/organic codes. In the current molecular code approach a sign (and meaning) is represented by the presence of a particular molecular species. There a countless alternatives, like a particular concentration interval of a particular species. (2) A second closely related question is: Can the molecular code be actually been used and how useful is it? If an ubiquitous species, like water, is a sign (or meaning), the code is obviously not usable. In another case, if the context (code-maker) cannot be generated and maintained by a system the code cannot be realized (implemented) and thus not been used. As a potential solution, this contribution explores an extension of the molecular code concept by including the formal concepts of self-maintenance and chemical organizations and the less formal concepts of autopoiesis and unity. The idea is to say that a code is usable, if a code-maker (context) exists that is contained in a chemical organization. Thus this requires a code-maker to be contained in a self-maintaining set Furthermore, we can demand that signs and meanings are not in the closure of this self-maintaining set. The concept is presented along various examples and its implications for analyzing molecular codes in real systems is discussed.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Code Biology is more fundamental than Quantum Biology

Peter R Wills
Department of Physics, Centre for Computational Evolution,
and Te Ao Marama Centre for Fundamental Enquiry,
University of Auckland, PB 92109, Auckland 1142, New Zealand
p.wills@auckland.ac.nz

Two main lines of reasoning attempt to link quantum mechanics to the foundations of biology. The first concerns molecular or collective phenomena which are observed in biological systems and which require specifically quantum effects, such as entanglement or tunnelling, for their explanation. Such phenomena now comprise a new, rapidly expanding field of enquiry which has been dubbed Quantum Biology. The second line of reasoning concerns the maintenance of structural and functional order in biological systems and is akin to answering the question "what is the essence of life?" with the retort "quantum information processing". There is a tendency for the first of these two ways of thinking to be conflated with the second, although Quantum Biology does not necessarily entail the notion that biological causation falls into a necessarily quantum rather than classical domain of nature. Fundamental theoretical biology has been approached from a quantum perspective for nearly a century with very little effect. The advent of the qubit as a unit of information has had no effect on the discipline of genetics simply because classical information theory and semi-classical theories of molecular behaviour give a demonstrably adequate causal explanation of genetic information, biological inheritance and the generation of phenotypic forms. However, even these classical theories fail to explain the emergence of self-sustaining control systems whereby polymer sequence information and signalling processes integrate the time evolution of molecular events in organisms on timescales ranging from microseconds to years and length scales ranging from nanometres to metres. The observed organic integration of coding and molecular events observed in biological systems, something which is completely arbitrary from the point of view of the underlying physics, does not arise especially from quantum coherence, rather from gratuitous formal relationships of "matching" between structures that first appear as accidental outcomes of molecular processes, thereafter giving opportunity for their own amplified growth. The quintessential role of coding in biology and its mode of emergence in complex molecular systems is illustrated by the self-organising process whereby a genetic code can evolve, beginning as a poorly defined binary mapping from one molecular alphabet (amino acids) to another (codons) and culminating in a functionally optimised form, which comprises a sophisticated "operating system" for the "naturally intelligent" control of the physical system in which it is embedded.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Implications & Teaching(s) of Code Biology

Robert Prinz
Friedberg (Hessen), Germany
robert-prinz@web.de

The presence and usage of codes is probably the most distinguishable and unique feature of living systems. All living entities use coded interactions at any level of their organization. There are no coded interactions in chemistry, physics, or geology. Yet, there are few if any definitions of life that take into account codes in the general life sciences. This shows, that although code biology and the field of biosemiotics are quite long-established disciplines, their implications for research and education still have an enormous potential to transform other disciplines. The presentation will address teaching principles and new research directions of code biology. Representation by means of cellular codes generates the cellular self - ”celf” - that is the basic building block of embodied wholeness. This wholeness or completeness is a setpoint realized and measured via cellular calculation and feedback loops. The author considers the celf as an equation and not a mathematical function. This working hypothesis has groundbreaking biological and medical implications ranging from cancer therapy to phantom pain treatment. Network analyses in biology have inherent weaknesses because they are based on one-to-one interactions between the analyzed molecules. Cellular information flow is facilitated by codes. Every code is composed of three parts: object, sign, and interpreter. This means, relations between two molecules are not necessarily based on direct chemical or physical interaction but on the mediation through an adapter, the sign. A new network topology and ontology is proposed by the author. Cellular components are assigned the roles object, sign, and interpreter according to Charles Sanders Peirce. This introduces a fundamentally different topology to networks compared to current approaches and opens new opportunities for the interpretation of cellular networks.
Today there is no simple and objective measure of complexity or even a satisfactory definition for the application of this term in the life sciences. Complexity of organisms or networks can be quantified objectively through the analysis of codes versus cellular building blocks. This measure allows comparisons of complexity between different entities, species, or cellular subsystems across biological realms. Also, quantification or estimation of complexity using codes can provide a measure of modifiability in synthetic biology: the more codes per parts, i.e. high code density, the higher the off-target rate will be during cellular engineering and molecular intervention. The definition of life must be revisited taking into account coded interactions. For example, a strict physical border or membrane is not necessary for the definition of life. A living entity is comprised of a set of codes ruling the interaction of its parts. Any interaction that is not compatible with this set of codes is not part of the system. This has also implications on the search for life in the universe. Finger- and footprints left by extraterrestrial life on other planets could be found through identification of codes and episemiotic signs in future space exploration.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Code biology and language: questioning the three systems'-view

Rasmus Gahrn-Andersen
University of Southern Denmark
Department of Language and Communication
rga@sdu.dk

According to Barbieri (2014), the history of evolution gives evidence of the existence of three different kinds of cognitive systems: The first, most basic kind involves mere perceptual capacities which that any living being including bacteria and other unicellular organisms can do (cf. Martin & Gordon 2001). The second kind of system pertains to eukaryotes having the neurophysiological complexity that not only permits them to perceive the world but also to interpret it. They do so as they "receive signals from the environment, transform them into mental images and perform mental operations" (Barbieri 2014: 143). Lastly, the third kind of system is the faculty of language which is human-specific and, specifically, initiated as a human infant makes use of "sounds to attract attention" and start mimicking the behaviors of others (Barbieri 2014: 144). This 'three systems'-view makes sense in evolutionary terms. Nevertheless, there are good reasons for questioning its suitability if it is taken as an ontological claim. Here, we need only take a closer look at one of the systems: language.

- First, it gives the impression that language builds on an interpretative dimension which can be potentially separated from linguistic activity. However, as Wittgenstein (2009) has shown, the basic attitude of 'grasping' something as something need not entail any mental reflectivity
- Second, it remains unclear how human linguistic activity is different from, say, the "interpretive" activity of birdsong which, as Farina (2017) shows, also builds on biological capacities for making "connections between emitting and receiving sounds" (Barbieri 2014: 144). Thus, what remains to be clarified is how human language (and the cultural complexity it has brought about) is different from other kinds of vocal sense-making.

This paper investigates the ontology of language in relation to the human ecology. It does so in an attempt at substantiating Barbieri's claim that language is qualitatively different from other cognitive phenomena.

References:

Barbieri, M. (2014). Code biology. London: Springer.
Farina, A. (2017). Ecoacoustic codes and ecological complexity. Biosystems. https://doi.org/10.1016/j.biosystems.2017.09.003
Martin, C. C. & Gordon, R. (2001). The Evolution of Perception. In: Cybernetics and Systems: An International Journal, 32: 393-409
Wittgenstein, L. (2009). Philosophical Investigations. Malden: Blackwell.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Quantum Biology and Rules of Probability in Long Genetic and Literary Texts

Sergey V. Petoukhov
Mechanical Engineering Research Institute of Russian Academy of Sciences.
101990, Moscow, M. Kharitonievskiy pereulok, 4, Russia,
spetoukhov@gmail.com

In historically the first work on quantum biology in 1932, P. Jordan wrote that the disclosure of the specifics of living organisms is associated with the discovery in the future of laws that are still overlooked. He claimed that «*life's missing laws were the rules of chance and probability (the indeterminism) of the quantum world that were somehow scaled up inside living organisms*». It is these laws of chance and probability, postulated by Jordan, that we purposefully seek and study in probabilistic characteristics of long DNA sequences of hydrogen bonds and nitrogenous bases. Some of our results are presented at the conference. Here one should explain the difference between biological and inanimate objects. Jordan pointed out that inanimate objects were governed by the average random motion of millions of particles, such that the motion of a single molecule has no influence whatsoever on the whole object. By contrast, in living cells the few molecules have a dictatorial influence, such that their quantum-level events are amplified to influence the entire organism. Jordan believed that living organisms were uniquely able to carry out this amplification. Jordan was convinced he could extend quantum indeterminism from the subatomic world to macroscopic biology. He even made a connection with free will by suggesting a link between quantum mechanics and psychology. The lecture presents our results of searching hidden rules of probabilities of elements in long sequences of hydrogen bonds 2 and 3 between complementary nitrogenous bases in DNA double helixes of different organisms, including a comparative analysis of these rules in complete sets of chromosomes in eukaryotes. Formalisms of quantum informatics with using the tensor product of vectors are used in modeling some of these results.
In addition, we show structural analogies between long genetic and literary texts taking into account known ideas of some authors (R. Jakobson, F. Jacob, et al) that human languages were formed not from an empty place but on the basis of the genetic language. These connections are shown using results of our comparative study of probabilities of elements in long DNA sequences of hydrogen bonds and in long Russian literary texts (novels by L. Tolstoy, F. Dostoevsky, A. Pushkin, etc.). These results add data to known analogies between genetic and literary languages. We continue these studies for literary texts in English, German, French and many other languages. The theme of Jung's archetypes is briefly discussed as well. In general, the results of our research confirm the adequacy of Jordan's ideas in the field of quantum biology.

References

Petoukhov, S. Structural Connections between Long Genetic and Literary Texsts. - *Preprints* 2018, 2018120142, https://www.preprints.org/manuscript/201812.0142/v1

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Evolution of Degeneracy in the Genetic Code

Simone Giannerini, Diego Gonzalez and Greta Goracci
University of Bologna, Department of Statistical Sciences
CNR/IMM Bologna Section
simone.giannerini@unibo.it

The origin and early evolution of the genetic code is one of the most challenging problems in evolutionary molecular biology. In this contribution we review a number of results on the origin and evolution of degeneracy in the genetic code. The same are based on a hypothesis on the symmetry of ancient adapter molecules and are consistent with a mathematical modelling of the genetic code first proposed in 2003 (Diego L. Gonzalez and Marcello Zanna, Sistema Naturae, 2003, Diego L. Gonzalez, Can the genetic code be mathematically modelled?, Medical Science Monitor, 10(04), HY 11-17, 2004). We are able to explain the origin of degeneracy in the early code, a genetic code ancestor of the LUCA's code that have been supposed to have the same degeneracy distribution that the extant Vertebrate Mitochondrial Genetic Code. We propose a pre-early code formed by codons and anti-codons of length four, the tesserae model. We discuss also new results regarding the conservation of the Rumer's anti-symmetry on evolutionary periods from the pre-early code to extant variants.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Languaging, writing and cultural codes

Stephen Cowley,
University of Southern Denmark (Slagelse campus)
cowley@sdu.dk

While languages can be described in terms of writing systems, they can also be traced to talking, reading, writing, singing, thinking etc. This *languaging* or "activity where wordings play a part" (Cowley, 2014) is the precursor of what is formalized around language systems. Indeed, the history of writing is a history of regularization and standardization: in a few thousand years, languaging has become ever more detached from bodies or autoglottic (Harris, 1996). Thus, while not intrinsically code-like, languaging nonetheless sustains artificial codes (that can be based on digits as well as various kinds of ideographic, phonetic and diagrammatic formalization). With writing, humans have opened up a 'theoretical culture' (Donald, 1991): they even have technologies that rely on codes to ensure that devices follow effective procedures. In the last 60+years, growth in computation has altered how human powers link direct or embodied understanding with literacy. All life forms use coding based on "a small set of arbitrary rules selected from a potentially unlimited number in order to ensure a specific correspondence between two independent worlds" (Barbieri, 2015). Humans apply these powers to create extended codes that allow for the use of texts, computers, phones and the Hubble (e.g. Clark, 1997; Giere, 2004). In human life, coded output – visible information – is often perceived as meaningful: extended systems couple with active looking, literacy, expertise etc. Indeed, Turing's pen-and-paper analogy (Wells, 2006) was implemented in the use of effective procedures (see Pinna, 2016) that sustain global markets, science and, of course, endless forms of entertainment etc. Within extended systems human agents (Giere, 2004) change rules and correspondence relations (e.g. interfaces, software and computer-use) to bring forth (knowledge of) black holes, markets and much besides. The results set off the eco-cultural effects that are a signature of wide coding. The claim can be readily defended. For example, with tools and trade, communities came to make normative use of visible traces: in time, skills with hand, eye and object gave rise to digits and inscriptions. Could these – like numbers and written words – arise from wide coding? There is a parallel with looking: vocal adaptors may attune to arbitrary material relations whose acoustics have functional, non-arbitrary properties. In Farina's terms (2018), communities co-construct 'soundscapes' with changing rules and correspondence relations. Yet, even if conceivably based in a soundscape, languaging also grants recursive power to articulatory gestures –we use wordings to repeat what we hear. Indeed, this too may be wide coding. Were this so, it would help explain how we came to inhabit social worlds where, in Maturana's (1983) terms, languaging enables us to observe and make observations.

References

Barbieri, M. (2015). Code biology. In M. Barbieri (Ed.) *Code Biology* (pp. 171-189). London: Springer.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge MA: MIT press.

Cowley, S. J. (2014). Linguistic embodiment and verbal constraints: human cognition and the scales of time. *Frontiers in psychology*, *5*, 1085.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press.

Farina, A. (2017). Ecoacoustic codes and ecological complexity. *Biosystems*. https://doi.org/10.1016/j.biosystems.2017.09.003

Giere, R. N. (2004). The problem of agency in scientific distributed cognitive systems. *Journal of Cognition and Culture*, *4*(3), 759-774.

Harris, R. (1989). How does writing restructure thought?. *Language & Communication*, *9*(2-3), 99-106.

Maturana, H. R. (1983). What is it to see?. *Archivos de Biologia y Medicina Experimentales*, *16*(3-4), 255-269.

Pinna, S. (2017). An extended approach to human computation. *Reti, saperi, linguaggi*, *4*(2), 289-302.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

Evolution Before life: how natural selection can take place in a dynamic population of lipid bodies

Wanderley Dantas dos Santos
State University of Maringa, Department of Biochemistry
Maringa, Parana, Brazil
wdsantos@uem.br

Thanks to its ability to form chains, in association with other elements (H, O, N, S) carbon can generate a virtually infinite diversity of organic structures with similar diversity of properties. In water, polar organic molecules can dissolve freely by establishing hydrogen bonds (H-bonds) with solvent molecules. In turn, non-polar compounds (lipids) cannot form H-bounds with water. Like this, on the interface with lipids water molecules self-assembly in a highly ordered structure. This high-energy structure dissipates when lipids group together, releasing part of molecules in the solvation shell to more entropic bonds. Like this, in water lipid bodies act as hubs entrapping all lipids and amphiphilic compounds that cross their way. When become excessively wide, lipid droplets become unstable and split in smaller ones. The size, splitting rate and other properties of the lipid population depend on the composition of lipids in the medium. As such a composition is not controlled by the lipid unity its properties cannot be transmitted to daughter-droplet's. But besides entrapping lipids, lipid aggregates also favor the occurrence of chemical reactions. Polar bounds performed by water molecules neutralize most electrostatic charges present in organic compounds. Therefore, the low dipolar constant found inside a lipid body allows the electric interactions that drive chemical reactions. In general, this catalysis is unspecific, but the specific chemicals present in lipid-body can provide some selectivity for those reactions. The mere presence of chiral compounds, for instance, can influence the asymmetry of an enantiomeric product. In addition, some small organic compounds are known to catalyze selective chemical reactions, a phenomenon known as organocatalysis. In certain conditions asymmetric organocatalysts can produce high enantioselectivity. Although the enantioselectivity is more studied due its application in organic synthesis, chemo- and regioselectivity have also been reported. In this way, the presence of compounds able to provide selectivity to reactions hosted in lipid bodies will not result in a random mixture of products, but in a biased group of products selected. As so, differently from unselective lipid bodies, selective lipid bodies can reproduce conserving chemical information despite the medium composition. Here, natural selection could take place selecting populations of lipid body populations with higher fitness in a medium (speed of splitting, stability, and so on). A way to completely conserve useful molecular information is an efficient resynthesis of the essential compounds. Recently, I proposed that life can be defined as chemical system capable to reproduce in its products the molecular information present in its components, a stricter way to mean resynthesis. In this way, prebiotic evolution in selective lipid bodies can rely on bases of the chemical evolution that gave origin to life, a process therefore older than life itself.

Sixth International Conference in Code Biology – Friedrichsdorf (Germany) 3-7 June 2019

The origin and role of DNA as a challenge for biosemiotics

Wim Beekman
Wageningen University and Research – Philosophy Group
wbeekman@filternet.nl

This case study is part of a philosophical study to elucidate the existence and role of information in living organisms, and their implications for current origin-of-life research. The comprehensive theoretical-biological project will among others, also pay attention to systems biology and emergence theory. What possibilities do different scientific approaches have for answering origin-of-life questions? Living entities have a few unique characteristics. They are things with complex internal dynamics, autonomous behaviour and interrelated parts. In sum, living things display a complex functional order (they are organized).
Living entities are self-organizing: they are striving to continuate a complex functional order. In this DNA plays a central role. For understanding the role of DNA in a cell not just a physical description but also an informational description is necessary.
The kind of information that DNA contains is multifaceted, since it is functional information. Its diversity is calculable (e.g. Shannon entropy). But more important is its biological specificity. By a process of decoding digital information in DNA has impact on analogue processes in the cell. Most common in explaining living entities are mechanistic approaches. 'Mechanism' understood as causal mechanism. Systems biology entails a mechanistic approach. Organicists conceive life as a phenomenon in its own right. Both approaches (mechanism and organicism) have their theoretical and practical strengths and weaknesses.
A diversity of methods is necessary to explain the role of DNA in living organisms: functional and reductive. The informational character of DNA makes biosemiotics indispensable. Mechanistic and organicist approaches in the field of biosemiotics have a complementary role. They interrogate each other when it comes to an understanding of the origination of DNA. These considerations will result in a preliminary conclusion regarding the state of the research on the existence and role of information in living entities.