Ninth International Conference in Code Biology

Guimarães (Portugal) 6 – 10 June 2023

ABSTRACTS

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Repurposing as a final step in the evolution of codes

Ádám Kun[1,2,3]
[1]Department of Plant Systematics, Ecology and Theoretical Biology, Institute of Biology, EötvösLoránd University, Pázmány Pétersérány 1/C, Budapest, H-1117Hungary
[2]Institute of Evolution, Centre for Ecological Research, Konkoly-Thege M. út 29-33 Budapest, H-1121 Hungary
[3]Parmenides Centre for the Conceptual Foundation of Science, Parmenides Foundation, Hindenburgstr. 15, D-82343 Pöcking, Germany

The evolution of codes of life goes through stages proposed by Marcello Barbieri: (1) beginning; (2) reducing ambiguity; (3) optimization; (4) major transition; and (5) conservation. This fits the evolution of the genetic code quite well. The genetic code, because its ubiquitous role in a living cell stuck at the conservation phase (cf. “frozen accident”). Other mappings on the other hand, might have gone on to evolve and change. From the evolutionary point of view, codes of life are powerful tools as the adapter can change to mapping. While the realms thus connected might not change, the exact mapping does. If the coding changes, then the same input can lead to a different output or a different input can produce the same input. We observe, for example in development of multiceullular organisms, that the key molecules and genes involved change very little through evolution, but the resulting forms produced do. This is achieved by often subtle differences in the targeting, tissue specificity, timing, etc. of the mapping. Consequently, I propose that the series of stages in the evolution of a code of life extended. Codes emerge (beginning), their ambiguity is reduced (evolution) and the mapping is optimized. The next step is conservation for a given purpose, but the potential for ambiguity is not eliminated. Then the code could later be used for other purposes. This phase might be called repurposing.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Revisiting of "The Multiple Codes of Nucleotide Sequences"

Alexander Bolshoy
Department of Evolutionary and Environmental Biology
University of Haifa, Haifa, ISRAEL 3498838
bolshoy@evo.haifa.ac.il

In 1989 E. N. Trifonov published his highly-cited article „The Multiple Codes of Nucleotide Sequences". In this pioneer paper Trifonov introduces his vision of the term "sequence code": "Code is a sequence pattern instructive for one or another specific molecular (multimolecular) interaction or process". As a matter of fact, these "Trifonov's codes", discussed in several publications of Trifonov and his coworkers, are of very various nature. In some cases, authors deal with instructions (chromatin code, translation framing code, protein folding code), in other cases Trifonov and others are talking about polysemanticity of a string (overlapping codes). In some publications we deal with rather a method for detecting codes (contrast words), in other cases we seem to be closed to what semiotically a code really is: a pattern with a given syntax coding for a given content. Here are the codes presented in Trifonov's paper: Translation Framing Code; Chromatin Code; Shape Code; Loop Code. He also discusses "overlapping and degeneracy of the codes" and hidden meaning of tandemly repeated sequences. We would try to retell contents of this influential paper following two major criteria: a) we would try taking into account that our reader is not an expert in Molecular Biology and she/he is glad to get our 30-years-after linguistic/semiotic retelling the tale of sequence code multiplicity; b) research of genetic sequence has never stopped and we would like to inform our reader regarding modern widespread opinions in the field. One of the aims of this talk is to show how a universalistic view of „genetic codes", often introduced into molecular biology through physics, has been crumbling under the onslaught of evidence from various branches of molecular biology over the past thirty years of molecular biology research. We will try to draw conclusions from this collapse of universalistic concepts in molecular biology. Further we would like to discuss what makes a „point-of-view" paper inspiring and widely cited because in spite of bringing wrong or non-convincing examples of DNA codes in that Trifonov's famous paper, it has become soundly influential.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Possible ways of origin of codes in the RNA World

Alexander Spirov
The Institute of Scientific Information for Social Sciences
Russian Academy of Sciences, Moscow, Russia
alexander.spirov@gmail.com

The concept of the RNA world postulates the possibility of random occurrence in the primordial soup of RNA molecules (ribozymes-replicases) capable of synthesizing their copies. Such a self-replicating ribozyme can multiply indefinitely, mutate, and inherit mutations. Accordingly, it is capable of demonstrating simple scenarios of molecular evolution (by beating the competition of replicators). The subsequent steps of prebiotic evolution should involve other ribozymes in the replication by such a replicator for more efficient (and more competitive) processes in the RNA World. What is important, the number of unverified hypotheses for the beginning of the RNA World is small. So far, experimenters have not been able to synthesize ribozymes capable of reproducing the beginning of the RNA World in an experiment. However, impressive advances in the synthetic biology of functional RNA molecules incline us to the theoretical possibility of the RNA World. At the same time, it should be borne in mind that alternative hypotheses of prebiotic evolution (such as the Protein world) are much less substantiated experimentally, and even theoretically. Therefore, it is the concept of the RNA World that it is reasonable to develop further. The big problem here is the scenarios by which protein enzymes and their encoding machinery emerge. In this communication, we will consider the most reasonable foundations and the first steps of the RNA World on the way to the World of Nucleic Acids/Proteins with a code (for example, the RNA-peptide world).

We believe that the general ability of functional RNA molecules to recognize signal sequences on other RNA molecules served as the foundation for evolution towards macromolecular codes. In modern biology, mention should be made of ribozymes that catalyze trans-splicing and trans-cleaving. The T-box riboswitches should also be included here. On a very long journey from signal sequences to codes, a significant milestone was believed to be the appearance of proto-transfer RNAs. The next major leap is the emergence of proto-aminoacyl-tRNA synthetases. And only after that we can expect the appearance of proto-ribosomes. At first, they could well have been compound complexes of ribozymes (without the participation of proteins). Thus, the closer we get in our consideration to the stage of the emergence of functional proteins and biological code, the less clear the concept becomes. At the same time, the experimental grounds for these hypotheses narrow down. However, it is the biological code that underlies the modern molecular machinery of life. Therefore, it can be argued that it is the code that is critical for the emergence of the kind of life that we observe on Earth today. It is quite remarkable that it is proteins that are able to read (recognize and interpret) nucleic acid sequences (signal sequences ranging from triplets to a couple of dozen bases). The ability of RNA (and DNA) molecules to recognize sequences in polypeptides is very, very limited. We are convinced that it is this asymmetry in the efficiency of reading signals that largely determined the evolution from the RNA World to worlds based on codes.

The study is supported by the Russian Science Foundation grant 22-18-00383.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Betwixt and Between

Minding the Brain

Anna Aragno

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

The demand for continuity has, over large tracts of science, proved itself to possess true prophetic power. *W. James, 1890, 97*

Out of the swamp of the reticular formation the cortex arose, like a sinful orchid, beautiful and guilty. *P. Yakovlev, in Konner 1982, 68*

The basic emotions are natural kinds that have specifiable neural substrates within the mammalian brain. If we do not come to terms with such foundation principles, we will have impoverished views of psychological and cultural complexities that ultimately arise from emotional learning. *J. Panksepp, 2018, 225*

The three Barbierian macro-evolutionary stages are traced in broad outlines marked by distinct manifest-shifts in semiotic advances preceded by long periods in which these developments were evolving. Within the first phase-transition between molecular/organic coding processes and the neural codes, I ask: What came before? What is pushing evolutionary processes in the human nervous systems? and if it is new coding forms, how did they develop, from where?

The second transition, from neural processes to natural-signal communications, I hypothesize, results as a consequence of affect-expressions at the genesis of subjectivity and 'mind.' The transition to iconic-signification, as in dreams, reveals the key role embodied-expressive signals have in developing human sign and representational capacities via expanded cortical interconnections unique to our species. The third transition addresses steps from iconic-signification to gestural/vocal-signs to a linguistic *system,* which universally becomes cultural coin in human communication.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Healing with psychedelics: Re-coding a number of codes

Candice Shelby
University of Colorado, Denver, CO 80217-3364
Candice.Shelby@ucdenver.edu

A major area of research in mental health treatment re-emerging after lying dormant for over 30 years, and one which is assuming priority in a number of important psychology and medical departments around the world involves the use of psychedelics. Psychedelic-assisted psychotherapy, and in particular the use of ketamine, psilocybin, and MDMA, has received increasing financial, academic, and social support in the past decade. Very recently, a second, Phase III confirmatory study showing the value of the use of MDMA in treatment of PTSD has been completed with positive safety and effectiveness results, setting the stage for submission of a New Drug application to the U.S.'s FDA. Another $2 million grant was just awarded to researchers of the University of Colorado Denver and New York University to replicate 3 other studies focused on the effectiveness of psilocybin administration in a psychotherapeutic setting for persons with end-stage cancer. The previous psilocybin studies, conducted at Johns Hopkins University, Stanford University, and the University of Colorado Boulder had found this therapy significantly more effective and safer than utilization of antidepressants for dealing with the psychological issues affecting this population.

This paper proposes to show that these substances operate by affecting natural biological coding processes at a number of levels, and in a way that does not simply treat symptoms, or require ongoing ingestion of medicines, but rather in a manner that seems to effect permanent healing (or at least semi-permanent; the patients were followed for 2 years after treatment with ongoing positive effects).For instance, classical psychedelics operate on neural codes. All of them act as agonists on serotonin 2(5-HT_{2A}) receptors. Seratonergic neurons project into virtually all areas of the cortex and higher areas of the brain, and they are involved in all kinds of emotions, from rage to hunger to lust to depression and mood more generally. Seratonin not as a neurotransmitter itself, but as a modulator of the major neurotransmitters such as GABA and glutamate, signaling neurons to enhance or inhibit their production, release, and/or re-uptake of these substances, which in turn, acting in various neural circuits, affects perception, attention, and, ultimately, through a number of other coding processes, significance and meaning.

Psychological changes in patients brought about by therapeutic doses of psychedelics seems to involve inactivating sub-routines that become embedded in our neural processing through the habitual firing of certain networks, which causes them, through yet another form of neural signaling, to act as attractor wells whenever any associated networks become active. When one particular neural network, referred to as the default mode network, a collection of neurons responsible for keeping these subroutines active, is suppressed, multiple other associations become active, which has the effect of quelling automatic responses and anticipations, so that the *meanings* of physical inputs, memories, and even of experience itself are open to change. Neural imaging shows that the effect of psychedelics is to diminish the inhibition of connections all over the brain, resulting in brain activity more characteristic of a 4 year-old than of an adult. Longstanding codings of things such as situations associated with substance abuse or trauma responses, and even death itself, can be altered through the use of these substances in well-defined therapeutic settings, with the result of long-term or permanent relief from symptoms of depression and PTSD. Philosophically, this alters the meaning of treatment for mental *health*.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Is meaning commensurable in scientific theories? From arbitrariness to non-nomological relations in meaning-making

Claudio J. Rodríguez H.
Palacký University in Olomouc, Czechia
claudiojrodriguezh@gmail.com

For scientific projects to deal with something as diffuse as meaning-making there are multiple hurdles to solve, starting with the validity of *meaning* as a specific phenomenon to be represented scientifically. Modulating the concept of meaning into a comprehensible phenomenon across different scales of validity---as a top-to-bottom approach in the sense of applying *meaning* to areas where it is not commonplace---requires being able to both differentiate its expression in subjective systems (as, for instance, individuals with the faculty of language) and as a biological principle (as something that happens at non-subjective or non-interpretive levels). In this presentation we will examine whether the latter sense of meaning can be somewhat commensurable with the former, and propose a philosophical change of gears in regards to the way we express the issues of meaning as arbitrary versus the non-nomological relations witnessed in accounts of biological meaning, making this the more accurate way to invoke the secondary sense of meaning without marring it with issues of subjectivity as it happens in the first sense.
Keywords: biological meaning; scientific theories; natural laws; downwards modeling

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

The Science of Codes and the Study of Signs: Two Contemporary Attempts at Understanding the Biological Nature of Meaning

Donald Favareau
National University of Singapore
favareau@gmail.com

For over 40 years, embryologist and theoretical biologist Marcello Barbieri has been developing a conceptual framework for understanding biological meaning based upon the analysis of the cell's internal organic codes. Originally christening this project 'Semantic Biology' (Barbieri 1985), Barbieri's physiological perspective on the code relationship differs drastically from of the Peirce-inspired sign theory that grounds Biosemiotics. Not surprisingly then, Barbieri rejects many of Biosemiotics' foundational assumptions (Barbieri 2014, 2019), while Barbieri's more fully-formulated version of Semantic Biology, now called Code Biology (Barbieri 2003, 2015), has likewise been met with some resistance from Biosemiotic theorists (Gare 2021, Markos 2013, 2019; Kull 2020). Yet for at over a decade, the two projects joined forces in the attempt to find common ground, and to work together to develop a community of researchers to pursue the still far too under-examined questions about the origin, nature and role of signs and codes in the organization and interactions of living systems. Many productive and long-lasting results came from this collaboration – including, but not limited to, the creation of book series, journals, international societies and annual conferences in both Code Biology and Biosemiotics. Too, the fact that the initial 'joining of forces' resulted in the formation of two now somewhat more independently thriving projects is not, in my opinion, a sign of failure, but of intellectual advancement and success, with each group now pursuing the investigation of meaning in the ways informed by the surfacing of the key points of disagreement between them over the course of their time working together (Favareau 2010).

For where Biosemiotics emerges from a tradition within theoretical biology that is organicist, qualitative and oft-times philosophically motivated, Code Biology is proudly mechanist and molecular and thorough-goingly empirical. The former, one might say, conceptualizes "the *biological nature* of meaning" from the more general to the more specific, while the latter conceptualizes "the nature of *biological meaning*" from the more specific to the more general. These are clearly two somewhat different projects from the get-go – yet when put in dialogue with one another, their interesting overlaps and critical distinctions offer much future food for thought. In this short talk, I would like to surface what I consider to be some of the more interesting and potentially productive nuances that arise when counter-posing key terms critical to each field, such as "natural convention" versus "habit"; "adaptor" versus "interpretant", and, of course, the relationship between "signs" and "codes." Knowledge of *both* Code Biology and Biosemiotics, I will argue, can only expand the understanding of anyone interested in the distinctive nature of living being.

References

Barbieri, Marcello (1985). *The Semantic Theory of Evolution.* London: Harwood Academic.
Barbieri, Marcello. (2003).*The Organic Codes. An Introduction to Semantic Biology*.Cambridge, UK: Cambridge University Press.
Barbieri, Marcello (2014). From Biosemiotics to Code Biology. *Biological Theory*, 9(2), 239-249.
Barbieri, Marcello. (2015). *Code Biology. A New Science of Life.* Dordrecht: Springer.
Favareau, Donald. (2010). *Essential Reading in Biosemiotics*. Dordrecht: Springer.
Gare, Arran. (2021). Code Biology and the problem of emergence.*Biosystems*208: 33-47.
Kull, Kalevi. (2020). Codes: Necessary, but not sufficient for meaning-making. *ConstructivistFoundations* 15.2: 137–139.
Markoš, Anton &Cvrckova, Fatima. (2013). The meaning(s) of information, code … andmeaning. *Biosemiotics* 6 (1) 37-51, pp 26-47.

Circular mixed sets

Elena Fimmel
(joint work with Christian Michel and Lutz Strüngmann)
Institute for Mathematical Biology, Faculty of Computer Science,
Mannheim University of Applied Sciences, 68163 Mannheim, Germany
e.fimmel@hs-mannheim.de

In this talk, the new mathematical concept of circular mixed sets of words over an arbitrary finite alphabet will be introduced. These circular mixed sets may not be codes in the classical sense and hence allow a higher amount of information to be encoded. After describing their basic properties, a recent graph theoretical approach for circularity will be generalized and applied to distinguish codes from sets (i.e. non-codes). Moreover, several methods are given to construct circular mixed sets. Finally, this approach allows us to propose a new evolution model of the present genetic code that could have evolved from a dinucleotide world to a trinucleotide world via circular mixed sets of dinucleotides and trinucleotides.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Can Code Biology Help Enhance Biolinguistics?

Fahad Rashed Al-Mutairi
Essex University (UK) and College of Education (Kuwait)
fahad.rashed@gmail.com

Biolinguistics, understood in the narrow sense as an offshoot of Chomskyan linguistics, is busy reshaping the results of linguistic theory in an effort to relate them to biology, physics and other advanced sciences. For this effort not to degenerate into loose correlations and vague metaphors, more seems to be required than just refining theoretical concepts and syntactic operations. What seems to be needed for a successful integration is a call to rethink some of the very foundations of Chomsky's linguistic framework. This call is somewhat implicit in Barbieri's admirable attempt to reach out to the field of linguistics, particularly through his semantic theory of language (Barbieri 2020).

Barbieri's semantic theory of language is embedded within his broader ideas concerning the semantic theory of evolution (1985), and this latter theory has later been replaced by Code Biology (Barbieri 2003, 2015). Code Biology, in the form of a semantic theory of language, deals with some of the fundamental questions pertaining to the nature of human language, its origin and evolution, and my aim here is to explore and examine the extent to which Code biology can shed light on the problems facing Biolinguistics. More specifically, since '*the code view of language*' contrasts sharply with '*the syntax view of language*', my strategy would be to assume the plausibility of the core hypotheses of Barbieri's semantic theory of language and then consider to what extent they are able to accommodate the results of linguistic theory while dispensing with some of its principles.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

The Quorum-Sensing Codes

Jan-Hendrik S. Hofmeyr
Department of Biochemistry, University of Stellenbosch
Private Bag X1, Matieland 7602, South Africa
jhsh@sun.ac.za

Intraspecies, intragenera and interspecies quorum-sensing-mediated cell-cell communication and synchronous group behaviours are not only the norm in the bacterial world, but also among eukaryotes and viruses. One example is the human host using quorum sensing to defend itself against bacterial invaders by harnessing its microbiome. Another example is a bacteriophage that monitors bacterial quorum sensing and kills the host bacterial cells only at high cell density. The quorum-sensing mechanisms are similar throughout: cells produce a signal molecule called an autoinducer and either an intracellular or a membrane-bound receptor for the autoinducer. The intra and extracellular autoinducer concentration increases with cell density until it reaches a threshold concentration where binding to the receptor induces a change in gene expression that activates a physiological activity or a change in the species composition of the community. Gram-positive and Gram-negative bacteria use quorum-sensing to regulate a wide range of physiological processes, such as symbiosis, virulence, competence, conjugation, antibiotic production, motility, sporulation, and biofilm formation. In general, Gram-negative bacteria use acylated homoserine lactones as autoinducers, whereas Gram-positive bacteria use oligopeptides. Although the chemical signals, the signal relay cascades, and the target genes controlled by bacterial quorum sensing systems differ, they all allow the bacteria to communicate with one another to coordinate the gene expression, and therefore the behaviour, of the whole community.

Quorum-sensing systems embody true organic codes, similar in nature to some signal transduction codes, but different in that they are community threshold codes where fixed but arbitrary rules govern communal behaviour. The autoinducers are the organic signs, the autoinducer receptors the adaptors, and the communal physiological responses the meanings. The LuxI/LuxR bioluminescence system of the marine bacterium *Vibrio fischeri* is the classical example of bacterial quorum sensing. It consists of two regulatory proteins called LuxI and LuxR. LuxI is the synthase enzyme that produces the autoinducer HSL (N-(3-oxohexanoyl)-homoserine lactone). HSL diffuses freely across the cell membrane so that its intracellular and extracellular concentrations are the same. As the *V. fischeri* culture grows, HSL accumulates to a threshold concentration somewhere between 1–10 µg/ml, which is sufficient for detection and binding by the cytoplasmic LuxR receptor protein. Interaction of LuxR with HSL exposes the LuxR DNA-binding domain. This allows LuxR to bind the *luxICDABE* promoter and activate its transcription, resulting in an exponential increase in both autoinducer production and light emission.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Life and death, or the dance of symbolic codes

João Carlos Major
International Academy of Analytical Psychology
Braga, Portugal
jcmajor@mail.telepac.pt

With *Symbols of Transformation* (CW 5), the psychiatrist and psychotherapist Carl Gustav Jung notes that in the course of the development of the psyche, in its attempt to adapt to the world (consciously), are used metaphors and comparations that in itself establish a passage between existing codes, overcoming the typical and stereotyped choices of the species. The result has been an enlargement of the image of the world and, over time, a renewal of former codes. That is, what Jung discovers is a process of continuous transitions between symbolic codes, which has a breaking effect in the context of rationality and consciousness and which is determined by the fact that the 'canonical grid' of behavioral codes (biological archetypes or patterns of behavior), in its monotonous determinism, no longer can capture phenomena and experiences as guaranteed until then, due mutation and variability of the environment. It will be increasingly evident that man, in the description of himself and the world, is in the circumstance of having to face anomalies and problems that, when they reach, in quality and quantity, a dangerously critical threshold can only be overcome if they change the 'metaphors we live by,' finding richer and more exciting analogies, introducing new and more vivid metaphors or codes. And finally, it seems that life 'invents' the death of individuals to overcome the threat of *sameness*. In this realm, Code Biology offers a renewed interpretive grid for Psychology in general and Analytical Psychology in particular, expanding the comprehension of human phenomena.

Barbieri, M. (2003). *The Organic Codes: An Introduction to Semantic Biology.* Cambridge University Press: Cambridge.
Barbieri, M. (2008). *The Codes of Life: The Rules of Macroevolution.* Springer: New York.
Cowley, J.C et al. (2010). *Signifying Bodies: Biosemiosis, Interaction, and Health.* The Faculty of Philosophy of Braga – Portuguese Catholic University: Braga.
Gazzaniga, M. (2018). *The Consciousness Instinct: Unraveling the Mystery of How the Brain Makes de Mind.* Farrar, Staruss and Giroux: New York.
Jung, C.G. (1956). *Symbols of Transformation - The Collected Works of C.G. Jung*, Vol. 5. Princeton University Press: Princeton, N.J.
Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By.* The University of Chicago Press: Chicago & London.
Neumann-Held, E. & Rehmann-Sutter, C. (eds.) (2006). *Re-Reading the Molecular Paradigm.* Duke University Press: Durhan & London.
Pieri, P.F. (2003). *Introduzione a Jung.* Editori Laterza: Bari.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

New Sequence Comparison Algorithm for Nucleotide and Amino Acid Evolutionary Relationships Based on Klein Four-Group (K4A)

Krunoslav Brčić-Kostić, Nikola Štambuk* and Paško Konjevoda
Ruđer Bošković Institute,
Bijenička cesta 54, 10000 Zagreb, Croatia
*stambuk@irb.hr

Phylogenetics is the study of ancestral relationships among biological species. Such sequence analyses are often represented as phylogenetic trees. The branching pattern of each tree and its topology reflect the evolutionary relatedness between analysed sequences. We present a Klein 4-group algorithm (K4A) for the analysis of relational databases of nucleotide and amino acid sequences. Relational databases are based on the relational model of data. This model of the Standard Genetic Code table is normalized and made up of nucleotides and related amino acids organized into four tables. Konjevoda and Štambuk (BioSystems 210:104529, 2021) showed that Klein four-group set of operators consists of: identity **e** (U/T), and three elements—**a** = transition (C), **b** = transversion (G) and **c** = complementarity, i.e. transition-transversion (A). We generated three K4A distance matrices based on the operator positions: 1. column positions, 2. row positions and 3. Cayley table positions. The performance of K4A was tested on a dataset of RecA proteins in bacteria, eukaryotes (Rad51 homolog) and archaea (RadA homolog). RecA and its functional homologs are found in all species. They are essential for the repair and maintenance of DNA (Michael M Cox. Nat Rev Mol Cell Biol. 8:127-138, 2007). Consequently, they represent a good model for the study of evolutionary relationship of protein or nucleotide sequences. The ancestral relationship between sequences was correctly classified by K4A. Three distance matrices exhibited small variations among species, and overall results of tree classification were in agreement with the general patterns obtained by standard PAM and BLOSUM substitution matrices. Our results indicate that different K4A operator positions within the Cayley table reflect diverse aspects of the cluster analyses: 1. column positions of the table specify clustering of protein sequences based on physicochemical amino acid properties, 2. row positions cluster sequences according to the coevolution theory, and 3. Cayley table exerts the classification based on complementary codons and error-correcting code. According to Marcello Barbieri (BioSystems 181:11-19, 2019) during the evolution of a code there is a phase of optimization of system rules, the ambiguity of a code is eliminated, and the system starts producing specific components. K4A algorithm enables use of different genetic code table variants optimized for particular transitions in evolution and based on biological specificity. K4A, Jukes-Cantor and Kimura models of DNA evolution will be examined, compared and discussed.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Turing and von Neumann machines: completing the new mechanicism

Barbora Jurková and Lukáš Zámečník
Palacký University in Olomouc
lukas.zamencnik@upol.cz

Turing (1937) introduces a model of code that is followed by other pioneers of computing machines (such as Thomas Flowers, John Mauchly and J. Presper Eckert and others). One of them is John von Neumann, who defines the concept of optimal code in the context of the conception of EDVAC (von Neumann 1945). He later uses it to build on in his theoretical considerations of the universal constructor (von Neumann 1966). Von Neumann (1951/1963) further presents one of the first neural network models, in relation to the work of McCulloch and Pitts (1943), for both theoretical purposes (von Neumann probe) and practical applications (computer architecture of EDVAC). The aim of this paper is (1) to describe the differences between Turing's and von Neumann's conceptualizations of code and the mechanical computing model. Between von Neumann's abstract technical conception (von Neumann 1951/63 and 1966) and Turing's more concrete biochemical conception (Turing 1952). Furthermore, (2) we want to answer the question why these influential models of mechanisms (predominantly in computer science) have so far been ignored by philosophers of the new mechanicism (Stuart Glennan, Peter Machamer, Lindley Darden, and Carl Craver). We will show (3) that these classical models of machines are not only compatible with the new mechanicism, but moreover complement it, since they represent a completely separate type of model of mechanism, alongside producing (Glennan 2002), maintaining (Bechtel and Abrahamsen 2005) and underlying (Machamer, Craver, and Darden 2000). We will prove (4) that these types of machines and the codes they represent find applications in some approaches of code biologists (Igamberdiev and Brenner 2021, Fimmel and Rodin 2021, Strüngmann and Shelah 2021), in modelling of living systems. Which testifies that the von Neumann probe is not just a historical milestone (Barbieri 2011) in the concepts of codes (Prinz 2022). The final and main goal of our paper will be an attempt to relate von Neumann's and Turing's notion of mechanism to Barbieri's notion of extended mechanism (Barbieri 2015).

References:

BARBIERI, Marcello: *A mechanistic model ofmeaning*in *Biosemiotics*. 2011. s. 1-4.
BARBIERI, Marcello: *Code Biology, A New Science ofLife.* Springer. 2015.
BECHTEL, William; ABRAHAMSEN, Adele:*Explanation: A mechanistalternative* in*Studies in History and Philosophyof Science Part C: Studies in History and PhilosophyofBiological and BiomedicalSciences*, 2005, 36.2: 421-441.
FIMMEL, Elena; RODIN, Andrei:*Thefoundationsofmathematics and theoretical biology* in*Biosystems*, 2021, 205: 104416.
GLENNAN, Stuart:*Rethinkingmechanisticexplanation* inPhilosophyof science, 2002, 69.S3: S342-S353.
IGAMBERDIEV, Abir U.; BRENNER, Joseph E: *Mathematics in biological reality: The emergence of natural computation in livingsystems*in *Biosystems*. 2021.
MACHAMER, Peter; DARDEN, Lindley; CRAVER, Carl F:*Thinkingaboutmechanisms*in*Philosophyof science*, 2000, 67.1: 1-25.
MCCULLOCH, Warren S.; PITTS, Walter:*A logicalcalculusoftheideasimmanent in nervousactivity* in*The bulletin ofmathematicalbiophysics*, 1943, 5.4: 115-133.
PRINZ, Robert: *A list od BiologicalCodes*, 2022.
SHELAH, Saharon; STRÜNGMANN, Lutz*: Infinitecombinatorics in mathematical biology.* in *Biosystems*, 2021, 204: 104392.
TURING, Alan*: On ComputableNumbers, withanApplication to theEntscheidungsproblem* in*Proceedingsofthe London Mathematical Society* s2-42, no.1 (1937): 230–65.
TURING, Alan*: Thechemicalbasisofmorphogenesis*in*Bulletin ofmathematical biology*, 1990, 52.1: 153-197.
VON NEUMANN, John:*First Draft of a Report on the EDVAC*, 1945.
VON NEUMANN, John: *The General and LogicalTheoryofAutomata*. ReadattheHixon Symposium in September, 1948; published in 1951. Collected Works 5.288-328.
VON NEUMANN, John: *Collected Works: Volume 5: Design ofComputers, TheoryofAutomata and NumericalAnalysis*. Edited by Abraham H. Taub. Oxford: Pergamon Press, 1963.
VON NEUMANN, John: *TheoryofSelf-ReproducingAutomata*. Urbana: University of Illinois Press, 1966.

Tell the Cell – How Key Are The words?

Marcella Faria
Dactyl Foundation for the Arts & Humanities
almeidapradomarcella@gmail.com

An Analogic Dictionary or "Thesaurus" is a compendium that organizes the words from a logical stand point, it establishes degrees of similarities between them according to abstract categories. There are many formal tools that can help in the definition of such categories. The use of word clouds and ontologies, for example, allows for a meaning-based classification of terms, by drawing semantic relations between words in their context of use, the sets emerge and, again, can be named as key-words. In the attempt to set the grounds for a "TheCELLrus" we have categorized the words coined around the term Cell (Cell-words) into discrete sets which emerged during the very process of classification and as we came to realize, reflected the expansion of seminal terms' semantic fields and metaphorical range. The term "Cell" itself, for example, has progressively (but not linearly or exclusively) stood for life's: place, essence, minimal unity, origin, agency, singularity, information repository, potential, limit. Two evolutive narratives had emerged: the epistemological one, chronology of discoveries and theoretical models; and the ontological one, the nexus between the actual biological novelties through evolution and development (structures, properties, processes, forms) and their impact on the symbolic potential of words. At last I will discuss the idea of a "TELLsaurus", in which to interweave the manifold metaphorical nexus between "Cell" and "Life" (as expressed by key words) with the manifold expressive strategies used in literary craft to tell singular fates. Be it as seen by theory and criticism (epistemic dimension), or as actualized in themes and formal solutions in poems and novels.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

A biology with hundreds of codes

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

The idea that life requires not only energy and information but also meaning was proposed for the first time in '*The Semantic Theory of Evolution*' (Barbieri 1985). The new biology described in that book was called *Semantic Biology* (a biology with meaning) but later on that name was changed into *Code Biology* because it is primarily the existence of codes that reveals the existence of meaning in nature. But how many codes exist in living systems? In 2012, when the Society of Code Biology was founded, the number of codes in its database was 22. Ten years later, in 2022, that number had gone up to 237 and there was no sign that it had reached the limit. It must be underlined that the Code Biology database reported only codes that were published in peer-reviewed journals and this means that the existence of hundreds of biological codes in living systems is based on sound experimental evidence. This gives us a variety of problems, theoretical and practical. On the theoretical level it seems that biological codes do not account only for the great events of macroevolution but for many other characteristics of the living systems. On the practical level we have the problem of the definition and of the classification of the biological codes as well as the problem of understanding why they are so different from the familiar codes of culture.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Meta-codes: handling impredicative relations across multiple scales to give meaning to structural and functional elements within a biosemiotic process

Mario Giampietro
ICREA, ICTA, Universitat Autònoma Barcelona
mario.giampietro1@gmail.com

A code is *a set of rules that establishes a correspondence, or a mapping, between objects belonging to two logically independent sets*. The coding operation entails the selection of a finite set of arbitrary rules determining the mapping across two finite sets of objects. When dealing with the phenomenon of learning/adaptation across multiple levels of organization and scales, it is obvious that "the" problem is how to define, update and integrate the definition and operation of codes. In relation to this problem, I suggest the existence of at least two meta-codes allowing the continuous updating and operation of the localcodes used in a biosemiotic process.

(1) the metabolic code. From non-equilibrium thermodynamics we know that dissipative structures can preserve an identity distinct from their environment by establishing a specific structural couplingdescribed by a generic semantic mapping (Schroedinger, and Prigogine) as **+dSi ←→ -dSe**. The generation of a surplus of entropy (+dSi) associated with the activities of self-organization taking place inside any metabolic system must be compensated by a flux of negative entropy (-dSe) generated by the context. This relation is a meta-relation linking: (i) the internal metabolic characteristics of the system (determining its ***umwelt***!); (ii) the external characteristics of its ***admissible environment***. That is, given the +dSi of a rabbit we can identify grass as its -dSe. Whereas the +dSi of a fox entails that rabbits are the source of -dSe. The meta-code +dSi ←→ -dSe allows to tailor, depending on the definition of +dSi (provided by the genetic code), the nature of the structural coupling of the different components of complex adaptive systems across different levels of analysis – i.e. the individual cell, the individual tissue, the individual organism, the individual species in the ecological niche, the functional compartments of the ecosystem. That is, the meta-metabolic code must be valid simultaneously across all levels of organization.

(2) the holonic code. Complex adaptive systems are organized in holarchies in which structural types (e.g. cells, tissues, organs, individuals, households – level n) get meaning from the functional types to which they belong (e.g. tissues, organs, individuals, households, communities – level n+1). This requires that: (i) structural types (processes taking place on the tangible side determining the structural coupling - the upward causation); and (ii) functional types (processes taking place on the notional side defining the functional coupling/purposes – the downward causation) have to be compatible in order to get a coarse graining - i.e. establishing tangible realizations of specific typologies of holon (the coupling of a structural-functional type). This is what I call **holonic coding**. The definition of types on the notional side (e.g. genotypes) allows the production of instances of types (e.g. phenotypes) on the tangible side. However, in order to be meaningful, the information used at any given level has to result compatible with the thermodynamic constraints determined by the metabolic code. That is, the genetic information of rabbits, cows, and goats must be compatible with the mass balance flow between herbivores, carnivores, primary producers and detritus feeders in the ecosystem.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

The wobble-effect and its Influence on genetic code variations optimized for the robustness against point mutations – a computer analysis

Markus Gumbel
(joined work with Elena Fimmel, Martin Starman and Lutz Strüngmann)
Center for Algorithmic and Mathematical Methods
in Medicine, Biology, and Biotechnology,
Mannheim University of Applied Sciences, 68163 Mannheim, Germany.
m.gumbel@hs-mannheim.de

Codon to amino acid assignments of the standard genetic code might help to minimize problems caused by point mutations. The robustness of a code against point mutations can be described with the help of a so-called conductance measure [1] – a weighted graph-based method. This presentation analyzes the influence of the wobble effect on genetic code tables and seeks for optimal robustness using an evolutionary optimization algorithm [2, 3] which optimizes the weights of the conductance graph. We demonstrate that the robustness is least influenced by mutations in the third position—like with the wobble effect. The results clearly demonstrate that point mutations in the first, and even more importantly, in the second base of a codon have a very large influence on the robustness of the genetic code. These results are put in context to single nucleotide variants (SNV) in coding sequences. The question is addressed which structure of a genetic code evolves from random code tables when the robustness is maximized. Our results illustrate that the evolving code tables are very close to the standard genetic code that the robustness against point mutations seems to be an important factor in the evolution of the standard genetic code.

[1] Błażej, P., Kowalski, D.R., Mackiewicz, D., Wnetrzak, M., Aloqalaa, D.A., Mackiewicz, P., 2018. The structure of the genetic code as an optimal graph clustering problem (preprint). bioRxiv. https://doi.org/10.1101/332478
[2] Fimmel, E., Gumbel, M., Starman, M., Strüngmann, L., 2021. Robustness against point mutations of genetic code extensions under consideration of wobble-like effects. Biosystems 208, 104485. https://doi.org/10.1016/j.biosystems.2021.104485
[3] E. Fimmel, M. Gumbel, M. Starman, L. Strüngmann: Computational Analysis of Genetic Code Variations Optimized for the Robustness against Point Mutations with Wobble-Like Effects. Life (2021), 11, 1338. https://doi.org/10.3390/life11121338

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Plurality of codes and coding types

Mikhail Ilyin
mikhaililyin48@gmail.com

The paper celebrates continuous extension of the list of biological codes Code Biology Database – A List of Biological Codes. The list includes 237 items as of December 2, 2022 and continues to expand. This fact alone poses a set of questions on how the amazing multitudes of biological and non-biological codes shape, reveal, and establish themselves, as well as how heterogeneous varieties of codes and coding match each other. The point of departure is the understanding of codes suggested by the definition in Code Biology Glossary – "A code is a set of rules that create a correspondence between two independent worlds" (entry Code - Code Biology_Glossary). This definition rests on three props or supports: rules, correspondence and two independent worlds. But surprisingly none of the three is in the glossary. My paper tries out to redress this shortfall to a degree. The paper first outlines a very diffused and overall notion of rules, controls and algorithms of all kinds. Then it proceeds to explore what the idiom *two independent worlds* may refer or stand for. Next the notion of correspondence comes to scrutiny. And finally, rules turn back, but this time very specific kinds of restraints, switches and algorithms that may serve to provide correspondence of the independent worlds in question. A major methodological expedience of the paper is the trick of primarying or getting back to basic principles and archetypical schemata. The paper specifically refers to what may be called archetypical distinction of the two complementary aspects of any integral phenomenon – its substantive matter-energy quanta and continua on the one hand and its formative information setups of structures and functions on the other. Each of the two worlds on its own provide complementarities of matter and energy, or information and meaning. Each of the two worlds is doomed to its own lot – the substantive one to entropy and "cooling down", while the formative one to negentropy and ordering up. Both 'worlds' are not *separate* entities but *distinct* (= independent of the code-biology definition) and *complementary* aspects of the integral phenomena. Actual and pragmatic worlds or phenomena co-exist and interact because correspondences between them and their complementary substantive and (in)formative copy-worlds provide affordances of all kinds for such interactions. The paper advances the idea that it is a complementary combination of phenomenal processes and their relevant mental (cognitive) schemata that results in emergent advancement of our evolving universe as we observe and know it. In other words,as soon as phenomena of various kinds interact with each other they turn into eco-evo-devo agencies, first quantal, then **physicochemical**, biologicaland finally human. Those agencies become operational because they copy or code each other (cf. entry **Copying and Coding** Code Biology_Glossary). Probably, on the levels of non-living phenomena one can add to copying also feedbacking, straightforward mirroring, autocatalyzing and catalyzing, or even elementary signal registering etc. The paper postulates an extensive domain of rules, techniques, regulations, algorithms, controls and switches based on their common family resemblance. With all the diversity of this domain its archetypical nucleus can be boiled down to an elementary act of divide-and-link by an interface. Such an interface may serve as a base for its open supplementing with all kind of functional and structural add-ons. The paper sums up its elaboration of the code-biology definitions of code and coding. This elaboration makes it possible to advance an over-all master interpretation of common properties ranging from the most elementary and simple ones to sophisticated and complex. The technique of simplex-complex transformations devised by the Center of advanced methods of social research and humanities helps to interpret and substantiate inherent integrity of coding / decoding phenomena as well their diversity and growing overabundance.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

How IUPAC Ambiguity Codes Specify Molecular Descriptors and Information Flow in Code Biology

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

The lecture is a systematic review of IUPAC ambiguity codes [1, 2] and their use in Code Biology. It is shown how to use a methodology of the nucleotide and amino acid encoding in order to extract accurate information on different properties of the biological systems. We also present statistical and data mining procedures that enable the use of IUPAC codes for the experimental measurements, validation and simplification of complex system modeling.

IUPAC ambiguity codes could be applied to:

1. encoding descriptive information of nucleotides, amino acids and proteins (e.g., of relative solvent accessibility, hydrophobic moment, intrinsic disorder, atom depth, etc.), and
2. systems modeling ranging from standard bioinformatic tools to classic evolutionary models (i.e., from Miyazawa-Jernigan statistical potential [3] to Kimura three-substitution-type model [4], respectively).

Coding theory properties and application of IUPAC ambiguity codes will be discussed, and emphasis will be given to their use for error-control coding and the relational data model [5, 6]. A particular attention will be paid to the interpretation of results in the context of the ambiguity-reduction theory by Marcello Barbieri [7]. Underlying mathematical, logical and semiotic concepts of interest will be presented and addressed.

References

[1] IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Abbreviations and symbols for nucleic acids, polynucleotides and their constituents. Recommendations 1970. Biochem J. 1970; 120:449-454.
[2] Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984. Biochem J. 1985; 229:281-286.
[3] Štambuk N, Konjevoda P, Turčić P, Kövér K, Kujundžić RN, Manojlović Z, Gabričević M. Genetic coding algorithm for sense and antisense peptide interactions. Biosystems. 2018; 164:199-216.
[4] Kimura M. Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci U S A. 1981;78:454-458.
[5] Štambuk N, Konjevoda P. Determining amino acid scores of the genetic code table: Complementarity, structure, function and evolution. Biosystems. 2020; 187:104026.
[6] Konjevoda P, Štambuk N. Relational model of the standard genetic code. Biosystems. 2021; 210:104529.
[7] Barbieri M. Evolution of the genetic code: The ambiguity-reduction theory. Biosystems. 2019; 185:104024.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Complexity, information and entropy, as the Moira's fate rulers of life

Enrique Farfán-Ugalde, J. A. Morales, O. Paredes
Bioengineering Translational Department
CUCEI, Guadalajara University
Guadalajara 44430, Mexico
omar.paredes@academicos.udg.mx

What minimum necessary elements determine biological systems, and how do they operate in emerging biological entities? We describe the complex organization, the capacity to manage entropy, and the biological information as essential characteristics of living entities. We explore life's information notion as an observable, proposing that informational ambiguity-reduction in organisms is closely related to their evolutionary process. In this regard, we suggest that evolution arises from a fluctuating convergence to a recurrent steady state - congruent to major evolutionary transitions - where an entropy-complexification balance yields such convergence. To reach a significant evolutionary shift, biological systems integrate complex structures embedded at different organizational levels, known as informational compartments. Starting from Barbieri's approach of syntax and semantics in code biology, we consider compartments as structural blocks - ranging from molecular to cultural scale - where a biological function results from an organic interpretation of organisms' information. We call such mapping -life code- the pragmatic dimension of code biology, where adapters make sense. On this basis, we conceive entropy as a physical cornerstone transversely coupled through biological machinery, assembling the processes described from code biology, informational theory, and complex systems theory. This new interdisciplinary code biology perspective explains the synergy among the biological descriptors of such a controversial phenomenon as life.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Ultra-conserved non-coding elements: Inferring coding principles from comparative genomics

Philipp Bucher
SIB | Swiss Institute of Bioinformatics
CH-1015 Lausanne, Switzerland
Philipp.Bucher@sib.swiss

Vertebrate ultra-conserved non-coding elements (UCNEs) have extraordinary properties. They are the most conserved DNA sequences on earth, showing 100% identity over hundreds of bp between human and mouse [1,2]. They co-occur in so-called genomic regulatory blocks comprising up to hundred or more individual elements [3,4].These blocks co-localize and bona fide control master regulatory genes, most of them known as transcription factors, and many of them linked to neural phenotypes. The working hypothesis is that UCNEs of the same block act in concert to fine-tune the expression of target genes with extreme precision. As of now, nothing is known about the codes, in which these elements are written. Searches for over-represented patterns have been inconclusive. The reader molecules have not been formally identified. Nevertheless, it is broadly speculated that UCNEs carry transcription factor binding sites, possibly arranged in overlapping fashion and intervening multiple times during development. A highly perplexing observation was that deletion of UCNEs resulted in viable, fertile mice with no obvious phenotype [5]. Proposed explanations for this paradox include redundancy and subtle phenotypes invisible in the lab. It took more than a decade until cellular anomalies in brain architecture could be demonstrated in mice lacking specific UCNEs [6]. The resolution of the "viable mice paradox" exemplifies how difficult it is to experimentally unravel the information contained in UCNEs. Progress towards understanding the molecular decoding mechanism seems even more elusive. Comparative genomics thus remains the only approach that could deliver observations exploitable for hypothesis testing. In my talk, I will present a computational framework based on evolutionary simulations, intended to be used to study the language, in which UCNEs are written. Language aspects amenable to this methodology include redundancy, positional inter-dependence, regular vs context-free grammars, modularity, and code overlap. Evolutionary parameters that may be varied include mutation and recombination rates, selection strength and mode, population size and genome ploidy. Comparative genomics observables serving as ground truth for hypothesis testing may consist of clade restrictions, synteny, length and percent identity of UCNEs, paralog patterns among others. Applying different evolutionary simulations to the "viable mice paradox" shows that high information density combined weak selection readily explains ultra-conservation in the absence of a strong phenotype while redundancy does not.

[1] Bejerano, G. et al.Ultraconserved elements in the human genome. Science. (2004).
[2] Dimitrieva, S. and Bucher P. UCNEbase--a database of ultraconserved non-coding elements and genomic regulatory blocks. Nucleic Acids Res. (2013).
[3] Kikuta, H. et al. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. Genome Res. (2007).
[4] Dimitrieva, S. and Bucher, P. Genomic context analysis reveals dense interaction network between vertebrate ultraconserved non-coding elements. Bioinformatics (2012).
[5] Ahituv, N. et al. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. Genome Res. (2007).
[6] Dickel, D.E. et al.Ultraconserved Enhancers Are Required for Normal Development. Cell (2018).

The Modularity Codes

Robert Prinz
Rechenkraft.net e.V., Marburg, Germany
robert-prinz@web.de

The modular organization of living organisms can be seen at the genetic, metabolic, cellular, and anatomic level, to name the most prominent (Prinz, 2022). The hallmark of modularity is the bundling of discrete components into structures that accomplish a certain function beyond the capacity of single building blocks. The involved components must pertain to that module more than they engage in connections or interactions with other components in the same cell or organism. Modularization therefore allows efficient distribution and segregation of tasks, while reducing unwanted interactions among too many components.
Today, it is still debated what enabled evolution of modularization itself. Overall, there are three very general levels to consider. First, the genetic level, where certain sets of genes comprise hereditary modules. Second, the developmental level, where sets of enhancers comprise morphological modules. And third, the metabolic level, where sets of enzymes comprise anabolic and catabolic modules. The overarching theme is how the involved players relate to each other, i.e., how molecular coherence enables coordination among the components and ensures their co-presence in space and time.
As shown by Marcello Barbieri, coded, tripartite, convention-like and arbitrary interactions permeate the living realm (Barbieri, 2006, 2019). Accepting this fact would implicate that in most modular structures consisting of multiple components, coded relations exist that connect the constituting components (1) amongst each other, and/or (2) to other components of the cell or organism. It will therefore be hypothesized that the number of coded relations among the components of a module is a measure of intra-modular coherence, while the number of coded relations involving components in and outside a module, determines the modularity of the system, or its granularity. In other words, codes are considered to be the basis of modular structures.
A comparative analysis of different known modules in terms of coded interactions among their components will verify or falsify this hypothesis. If true, the code distribution over an entire cell could then be used to identify yet unknown modules solely based on a certain regional or temporal code density over distinct sets of components. At the same time, codes would qualify as a source of modularity in the first place, due them being minimal (tripartite) modules on their own with a highly “connective potential”.

References

Barbieri, M. (2006). Life and semiosis: The real nature of information and meaning. *Semiotica*(158), 233. https://doi.org/https://doi.org/10.1515/SEM.2006.007
Barbieri, M. (2019). A general model on the origin of biological codes. *Biosystems*, *181*, 11-19. https://doi.org/https://doi.org/10.1016/j.biosystems.2019.04.010
Prinz, R. (2022). The modularity codes. *Biosystems*, *219*, 104735. https://doi.org/10.1016/j.biosystems.2022.104735

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

Biological codes, translation and bioceonoses

Sergey V. Chebanov
Department of Mathematical Linguistics
Philological Faculty of St. Petersburg State University
St Petersburg, Russia, 199034

s.chebanov@gmail.com

The most important events in the formation of Code Biology were the discovery of signs in nervous activity of animals (the end of the 19th - the beginning of the 20th centuries), the realization of their significance for ethology (the first half of the 20th century), the discovery of the genetic code and the establishment of its universality for all living organisms (1950-60s), revealing the diversity of semiotic means of living organisms (second half of the 20th century) and realization of the plurality of biological codes (beginning of the 21st century). As a result, semantics, expelled during the struggle against anthropomorphism, returned to biology. An important circumstance was the realization that a living organism is not a sign, but a self-reading text (K. Kull, 2000). However, the meaning and mode of existence of multiplicity of codes remain unclear. In this regard, there is reason to pay attention to the fact that organisms do not exist in isolation and not even as populations of one species, but as multi-species biocoenoses. In turn, biocoenoses are comparable to cultures. In the context under discussion, the semiotic interpretation of the nature of culture by Yuri Lotman is quite remarkable. Its essence is as follows. Each culture is a unity of many coexisting semiotic systems. These semiotic systems are mutually untranslatable. Nevertheless, there is a constant translation from the language of one semiotic system into the language of another one. Since these semiotic systems are untranslatable, the resulting translation turns out to be more or less non-equivalent. It is this non-equivalence that is the source of innovations that ensure the development of culture. In the considered aspect of comparing culture and biocoenosis, Lotman's understanding of the nature of culture makes it possible to understand the need for a multitude of semiotic systems inherent in biocoenoses of organisms. Then the biocoenosis, like culture, acts as a set of semiotic systems, both endosemiotics, functioning inside the body (genetic code, hormones, including tissue hormones, neurotransmitters, immunity factors), and exosemiotics between organisms (smell, pheromone, color, sound, gesture, posture, behavioral communication). At the same time, at every moment there is a translation of texts created by one code into texts created by another code (for example, genomes, as texts created by means of the genetic code, into cellular metabolism, created by means of the biosynthetic code, etc.). Such a translation is obviously ambiguous, and therefore non-equivalent, since it depends on the previous history of the cell, its functional state, the state of the whole organism, environmental conditions, etc. However, such non-equivalence can provide the adequacy — i.e. the optimality of the translation for the current state of an organism. On the basis of the developed point of view, the most different processes can be considered: epigenesis and its incomplete determination by the genotype, the state of the biocoenoses and its ambiguous relationship with the organisms composing it, in particular small or invasive species, different options of the succession of biocoenoses of the same type, etc.

Sponsored by grant 22-18-00383 (RSF).

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

The principle “like begets like” in algebra-matrix genetics and code biology

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

This material is devoted to analysis of emergent properties of the system of binary oppositions in the genetic code ensemble. The epochal model of the double helix of DNA by Watson and Crick showed that the multiple reproduction of genetic information on DNA strands uses the ancient principle “like begets like” based on the simple complementarity in pairs of nucleobases A and T, as well C and G. Each of these pairs contains a particular case of the binary oppositions “purine-pyrimidine”. But the system of DNA alphabets and genetic coding is much richer in types of binary oppositions, including molecular oppositions "amino-keto", "purines-pyrimidines", "strong and weak hydrogen bonds”, as well as "strong and weak roots of triplets" and "*n*-plets with strong and weak roots".

When studying the emergent properties of this system of binary oppositions in matrix form, using molecularly based binary numbering of components in algebra-genetic matrices, it is revealed that the entire system of genetic coding is built on the principle “like generates like”. This is manifested in algebraic correspondences of parts of genetic matrices when using the operation of binary-complementary replication: mutual replacements of binary numbers 0 and 1 in the binary numbering of parts in matrices. The named correspondence concerns the connections of these genetic matrices with algebras of multidimensional numbers in their matrix representation (in particular, a connection with the Poincaré conformal disk model of hyperbolic geometry). In the structure of genetic matrices, an analogy with complementary replication of DNA strands is presented.

Revealing the key genetic importance of the principle “like begets like” allows thinking about the general heritable meaning of this principle in genetically inherited physiological systems and multi-level interrelations of code biology. In particular, it concerns bio-morphological symmetries; the universal rules of stochastic organization of genomic DNAs; our visual perception with its optical system transmitting external images to the retina in complementary inverted and reduced forms, which are complementary replicated in the brain for decoding estimations [1-3]. It concerns also the existence of mirror neurons in the brain of human and animals, which are considered by many as participants of the cognitive activity, an origin of languages, automatic imitation, and other things connecting with code biology. The presented results give pieces of evidence that the systems of mirror neurons and DNAs complementary replications are not isolated parts of the organism, but they are included in a holistic bio-algebraic complex realizing the inherited principle “like begets like”.

References

1) Petoukhov S.V. The Principle "Like Begets Like" in Molecular and Algebraic-Matrix Genetics. *Preprints* 2022, 2022110528 (doi: 10.20944/preprints202211.0528.v2).
2) Petoukhov S.V. Binary oppositions, algebraic holography, and stochastic rules in genetic informatics. *Biosystems*, vol. 221, 104760, November 2022. https://doi.org/10.1016/j.biosystems.2022.104760
3) Petoukhov S.V. The stochastic organization of genomes and the doctrine of energy-information evolution based on bio-antenna arrays. *Biosystems*, 2022, 104712, ISSN 0303-2647 https://doi.org/10.1016/j.biosystems.2022.104712.

Ninth International Conference in Code Biology – Guimarães (Portugal) 6-10 June 2023

On the Coding Relevance of the Nucleotides from the phonological point of view

Suren Zolyan
Institute of Scientific Information on Social Sciences
Russian Academy of Sciences, Moscow,
surenzolyan@gmail.com

1. The only prominent molecular biologist who likened nucleotides to phonemes was Fr. Jacob (1977), probably influenced by his interlocutor, Roman Jakobson (Jakobson 1970, 438). Phoneme, unlike sound, is an element of the language system and is defined as a minimal set of features that distinguish it from other phonemes. Extrapolation of this methodology to the genetic code leads to a distinction between a nucleotide as a biochemical element and a nucleotide as an intrasystem abstract unit of the standard genetic code (SGC).The relevant differences between nucleotides can be represented as two pairs of binary oppositions based on two distinctive features: the number of carbon ring bases (two vs. one); the number of hydrogen bonds: three (G, C) or two (A, T / U). It is these differential features that can be considered the minimal units of SGC. Their ontology is determined by intra-system encoding and decoding operations.
2. The SGC may be represented as a superposition of two coding systems: a) quasi-triplet (32 codons can be represented as "doublet + comma", when the third position separates, but not distinguish triplets from each other); and b) semi-triplet - 30 codons are coded according to the principle: "doublet + purine" or "doublet + pyrimidine". The proper triplet coding is observed only in two codons encoding tryptophan and methionine when all three positions are relevant. One can find out the gradation of coding complexity – quasi-triplet (32 codons), semi-triplet (30 codons), proper triplet (one case - tryptophan), and context-dependent triplet (methionine/start-codon).
3. As an analogue of syntagmatic (pairing) and paradigmatic (substitution) relations, one may describe the process of genetic translation, reading and recognition. As it happens in phonological systems, there are strong and weak positions for sign distinction. Neutralization occurs in weak positions - in the so-called degenerate codons, the third position is irrelevant, and in 32 cases, can be filled with any nucleotide (neutralization of both features), and 30 – with the nucleotide of the same group (neutralization of the number of hydrogen bonds). The same asymmetry is observed in syntagmatic relations, manifested as the unreadability of the third position. This extrapolation allows us to offer a new interpretation of the wobbling phenomenon. The wobbling and super-wobbling are not a deviation from the regular pattern, but the different reading regimes: the first and second entries are regulated by more rigid rules (equal number of links + different number of rings) than the third.

References

Jacob F. The linguistic model in biology // Roman Jakobson. Echoes of his scholarship /. Eds. D. Armstrong, C.H. van Schooneveld. – Lisse: Peter de Ridder, 1977. – Pp.185-192

Jakobson, R. Linguistics in relation to other sciences. In Roman Jakobson, Selected Writings: Vol. 2:(1971). 655–696, The Hague — Paris: Mouton.

Physiological Engineering: A Code Biology Inspired Technology

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

The seminal article of Gordon M. Tomkins (1974) provides a convincing evolutive narrative on the origins of the *metabolic code* that mediates cell external stimuli with a metabolic responses through symbolic compoundsas cAMP, ppGpp, Ca, inositol phosphate and so on. Tomkins also speculates on how long-lasting hormones eventually substituted secondary messengers as physiological symbols between progressively distant sensory and receptor cells in larger and larger animals. It would not be a long shot to extend Tomkins' reasoning to plants, which then would allow us to talk in terms of *physiological codes*. Indeed, plants present compounds universally recognized as phytohormones, whose definition diverges of that of hormones mainly because plants have not a circulatory system nor glands. Some phytohormones as auxin, gibberellins, cytokinins, ABA, brassinosterols and so on have the biosignaling cascade already known in great detail. But plants also present a great number of compounds that perform signaling roles in a way that is not so clearly identifiable as phytohormones. Plants produce, collectively, more than *80 thousand known natural compounds*, which they use to cope with the challenging environment of their sessile way of life. Frequently, plants use this incredible array of compounds for communicate with pollinators, disseminators, symbionts, and inclusive with other plants, a process known as allelopathy. Studying the biochemical mechanisms of these physiological and ecophysiological processes of plant and plant cell communication, my group and I were able to develop a new technology – we dubbed *physiological engineering of plants* – to induce characteristics of interest in crops. For this, we applied known signaling compounds and also designed artificial semiochemicals that are then tested in greenhouse scale. The most promising active principles are than formulated for leaf spraying and tested in agronomical scale. This way, so far we were able to formulate agrochemicals capable of reprogramming plant physiology to produce: 1) sugarcane, soybean, maize and brachiaria crops with saccharification prompt biomass for cellulosic ethanol production (dos Santos and Buckeridge 2018; Bevilacqua et al. 2019; Ferro et al. 2020a and 2020b; Parizotto et al. 2020a and 2020b; Martarello et al. 2021; Mendes et al. 2022; dos Santos et al 2023); 2) hyperlignified crops with stems, leaves, fruits and grains more resistant to mechanical and biological stresses (Gonçalves 2022), and; 3) fast-growth tree seedlings for rapid urban afforestation and reforestation (Salatta et al. 2023).

References

Tomkins (1974). *Sci*, 189:760-763.DOI:10.1126/science.169570.
dos Santos and Buckeridge (2018).INPI, PI 1104756-9.
Bevilacqua et al. (2019).*Plant Physiol Biochem*, 142:275–282.DOI:10.1016/j.plaphy.2019.07.015.
Ferro et al. (2020a). *Process Biochem*, 90:131-138.DOI:10.1016/j.procbio.2019.11.024.
Ferro et al. (2020b).*Plant Biochem Biotech*, 29(3):484–493.DOI:10.1007/s13562-020-00561-0.
Parizotto et al. (2020).*Plant Physiol Biochem*, 151:421-428.DOI:10.1016/j.plaphy.2020.03.053.
Parizotto et al. (2021).*Plant Mol Biol Rep*, 39:179–191. DOI: 10.1016/j.plaphy.2020.03.053.
Martarello DCI, Tonete-Diniz DC et al. (2021).*Biomass Conv. Bioref*. DOI: 10.1007/s13399-021-01842-x.
Mendes et al. (2022).*Plant Physiol and Biochem,*179:12-19. DOI:10.1016/j.plaphy.2022.02.018.
dos Santos et al.(2023).*Biomass Bioen*,168:106684. DOI:10.1016/j.biombioe.2022.106684.
Gonçalves, DEG. (2022). Lignin induction on soybean crop under different edaphoclimatic conditions and phenological stages. MastersDissertation in Cell and Molecular Biology.
Salatta, FV. (2023). Promoting of growth and lignification in young tree plants and degraded pasture with signaling compounds. PhD thesis in Cell and Molecular Biology.