

## Second International Conference in Code Biology

Jena, 16-20 June 2015

### ABSTRACTS

- 1 Anna Aragno
- 2 Zachary Ardern
- 3 Marcello Barbieri
- 4 Gérard Battail
- 5 Andres Burgos
- 6 Han-liang Chang
- 7 Joachim De Beule
- 8 Peter Dittrich
- 9 Almo Farina
- 10 Elena Fimmel
- 11 Simone Giannerini
- 12 Louis Goldberg
- 13 Diego Gonzalez
- 14 Dennis Görlich
- 15 Markus Gumbel
- 16 Martin Hanczyc
- 17 Jannie Hofmeyr
- 18 Tim Ireland
- 19 Kristian Kraljic
- 20 Stefan Kühn
- 21 Ádám Kun
- 22 Barry McMullin
- 23 Chris Ottolenghi
- 24 Daniel Polani
- 25 Sonja Prohaska
- 26 Silke Sachse
- 27 Peter Stadler
- 28 Darko Stefanovic
- 29 Lutz Strüngmann
- 30 Eörs Szathmáry
- 31 Morten Tønnessen
- 32 Jochen Triesch

## Dream Codes, II

### Deciphering the Language of the deep Unconscious

Anna Aragno

After years garnering a reputation as outstanding research biologist, neuro-pathologist, medical theorist of the etiology of hysteria, even Cocaine experimenter, Freud now daringly applies his cumulative clinical observations and interpretive acumen to decoding the pictographic language of the common dream. “The Interpretation of Dreams,” (1900) emerges amidst a combination of wonder and ridicule from a Viennese medical community not uniformly swayed by Freud’s psychical turn. Working in his legendary “splendid isolation” Freud has produced a masterwork of scientific observation which, in chapters six and especially seven, presents a detailed analysis of the two tiered structure; the motive force; the formal properties and compositional grammar and syntax of a ‘primary process’ vocabulary, the language of the deepest unconscious in which the meanings of dreams are spun. He has arrived at these landmark observations, and *a scientific method of interpretation*, through the conceptual syntheses of the biogenetic law; the unconscious as phylogenetically archaic; the primacy of early experience; the power of repression, fixation, and regression, in sleep; and, most importantly, the sharply dichotomized primary (impulsive) and secondary (inhibitory) processes, as cognitive *principles* of mental functioning. Couched in his first *topographical* theory of mind, (Ucs. Pcs Cs), the cornerstone of psychoanalytic metatheory, with the “Interpretation of Dreams” Freud established himself as the fountainhead of a general dynamic psychology.

When listening clinically to verbal accounts of pictographic dream narratives we are hearing the expression of sensory-affective experiences still tied to perceptual impressions and *feelings* in the very process of being transformed into pre-cognized re-presentations. These iconic clusters, stretching from sensory-kinetic-emotive *physical* experiences, to memories, motives, fears, and desires; exhibiting proto-semiotic mechanisms of pre-linguistic tropes and cognition, obey laws that reveal a “psychic reality” that is the true subjective experience of the dreamer. With its metaphorical and manifest/latent structure, the dream is our MRI into the human unconscious, as valuable an investigative instrument into primitive forms of early human mentation as it is for uncovering individual psychical beliefs and states.

This presentation will cover in detail the mechanisms of the dream’s ‘primary process’ meaning-forms—the actual code of dreams— as laid out in chapter seven of the dream book including an eye toward a contemporary re-interpretation of some Freudian ideas based on updated integrations from an interdisciplinary palate of disciplines such as neuroscience, semiotics, linguistics, dialogics, and narrative theory.

## **Deterministic and Optimizing Processes in Theories of Genetic Code Origin**

Zachary Ardern  
School of Biological Sciences, University of Auckland  
[zard001@aucklanduni.ac.nz](mailto:zard001@aucklanduni.ac.nz)

In this paper I explore the relationships between chance, necessity and adaptation in different theories of the origin of the genetic code. Two important areas of research in the origin of the standard genetic code are stereochemical theories and work on various ways in which the codon table is 'optimised'.

I review the evidence for stereochemical theories, and the different kinds of optimality which have been observed in the genetic code. Literature in the two fields seems to have largely diverged, and there is a degree of tension between their basic premises - stereochemical mechanisms are deterministic and act independently of any functional results, while optimality implies some kind of adaptive process of optimisation. One finding which decreases the room for compatibility between the two branches of theory is the fact that stereochemical binding between codons or anticodons and their associated amino acids has predominantly been found for complex amino acids generally thought to have been added to the code after its initial stages. The concept of 'coevolution' of the code, constrained by the biochemical requirements of amino acid synthesis, is another deterministic process which has been invoked to explain the standard codon table. I also discuss how features of the code which are 'adaptive' in modern biological systems such as eukaryotic cells might relate to early forms of life, arguing that 'modern' life has made remarkable use of features present from early stages of the code.

In light of these issues, I challenge the assumption that the genetic code was built up through a pseudo-Darwinian gradualist mechanism involving searching vast areas of the information space of possible codon tables. Instead I argue that while many questions are left unsolved, the early development of the genetic code must have been constrained by many deterministic factors, which in hindsight appear aligned with adaptations which biological systems subsequently achieved.

## Ancestral, Ancient and Modern Genetic Code

Marcello Barbieri  
Dipartimento di Morfologia ed Embriologia  
Via Fossato di Mortara 64a, 44121 Ferrara, Italy  
[brr@unife.it](mailto:brr@unife.it)

The modern genetic code is a mapping between 64 codons carried by transfer-RNAs and 20 amino acids carried by 20 aminoacyl-tRNA-synthetases, each of which attaches one amino acid to one or more tRNAs. The synthetases are specific proteins that can be made only by an apparatus that already has a genetic code, and this gives us a classic *chicken-and-egg* paradox: how could the genetic code come into existence if it is implemented by proteins that can be produced only when the code already exists? The logical solution to this paradox is that the *modern* apparatus of protein synthesis was preceded by an *ancient* apparatus where the amino acids were attached to the transfer-RNAs not by proteins but by RNAs. The *modern genetic code* based on protein-synthetases, in short, was preceded by an *ancient genetic code* based on RNA-synthetases. This ancient code, in turn, evolved from previous codes that here are collectively referred to as *ancestral genetic codes*. The most primitive ribosomes were probably pieces of ribosomal-RNAs that were sticking amino acids together at random and were producing statistical proteins. The ancestral genetic codes, in other words, were *ambiguous* because a codon could code for many amino acids, and a sequence of codons was translated some time into a protein and some other time into a different protein. The modern genetic code, on the other hand, is *non-ambiguous* because every codon codes for one and only one amino acid, and evolved from an ancient code based on RNA-synthetases. This means that there have been three distinct phases in the history of the genetic code: (1) the origin of the ancestral genetic code, (2) the evolution from the ambiguous ancestral code to the ancient code, and (3) the evolution from the ancient to the modern genetic code. Our problem is to figure out, at least in principle, how could these three major transitions have taken place.

## Genomic error-correcting codes shape the living world

G rard Battail  
Retired from E.N.S.T., Paris  
email: gb@gerard.battail.name

Any material object is subjected to an erosion due to the chaotic behaviour of matter at the molecular scale. The second law of thermodynamics states indeed that its physical entropy, which measures how disordered it is, can but increase as time passes. Some sequence of symbols written on such an object, however, can escape degradation by means of an *error-correcting code*, a very sparse subset of all possible sequences to which the sequence to be protected must belong. Its elements are referred to as *words*. The actually written codeword can unambiguously be determined despite a number of symbol errors if the codewords are different enough, as a result of the code scarcity (i.e., its *redundancy*) and of an even distribution of its words. In the presence of a steadily increasing number of symbol errors, the sequence can be almost surely conserved provided its recovery is attempted after a short enough time interval and the recovered sequence is exactly rewritten, these operations being moreover regularly repeated. Instead of an unavoidable erosion in the absence of coding, doing so ensures the sequence conservation, except for a probability of failure which is very low if the code is efficient enough. In case of failure, a sequence markedly different from the original one results. As another codeword, it is conserved by the recovery process just as efficiently as the original one.

The DNA molecule is a material object at the molecular scale, so the permanence of the symbolic sequence it bears, the *genome*, demands that it belongs to an error-correcting code. That the living world is organized into discrete *species*, and not made of chimeras, directly results and makes a taxonomy possible. This 'main hypothesis', however, is not enough to account for the better conservation of very old parts of the genome like the *HOX* genes. It should be further assumed that the genomic error-correcting code is made of *nested component codes* which successively appeared during the geological ages. This 'subsidiary hypothesis' explains that the taxonomy is actually hierarchical. The genome conservation process by almost periodic recovery attempts (which are almost always successful) matches the existence of successive generations in the living world, which should then be interpreted as *regenerations*. The trend of evolution towards increasing complexity is another consequence of these assumptions, since information theory tells that lengthening a code can improve its error-correcting ability. The Darwinian selection operating on the error-correcting ability of genomes of increasing length then results in increasingly complex phenotypes.

The total number of extant and extinct species is very small as compared with the huge number of nucleotide combinations that even the smallest genome can assume. The living world as a whole thus matches the description given above of an error-correcting code: the genomes of actual species are very sparse within the set of all possible nucleotide combinations. Besides endowing any species with a unique label, a genome is also the recipe which controls the assembly of a phenotype by means of some syntax which provides the needed redundancy. The means by which living beings originate are thus identical to those which enable their conservation.

## Informational parasites in code evolution

Andres C. Burgos and Daniel Polani

[a.c.burgos2@herts.ac.uk](mailto:a.c.burgos2@herts.ac.uk)

[daniel.polani@gmail.com](mailto:daniel.polani@gmail.com)

We consider the problem of the evolution of a code within a structured population of agents. The agents try to maximise their information about their environment by “listening” to information from other agents in the population. Which agents “listen” to which other agents is determined by the structure of the population, which induces a “conversation graph” among agents. The traditional use of information-theoretic methods would assume that every agent knows how to “interpret” the information offered by other agents. However, central to this is the assumption that one “knows” which other agents one observes, and thus which code they use. In our model, however, we specifically preclude that: it is not a priori clear which other agents an agent is observing, and for understanding a code the agent either must know the identity of the other agent with which it is communicating, or the code must be universally interpretable. A universal code, however, introduces a vulnerability: a parasitic entity can take advantage of it [1]. Here, we investigate this problem. We consider a parasite to be an agent that tries to maximise its information about the environment, while, at the same time, inflicts damage in the mutual understanding of the rest of the agents of the population. After introducing a parasite in a structured population, we let the population respond to this new addition by globally optimising the mutual understanding of the agents via rearrangements of their conversation graph. This optimisation from a global perspective indeed selects the optimal strategy, namely the isolation of the parasite. However, this assumes a global, unconstrained update of the interactions between each of the agents. Now we consider, instead, a constrained evolution of the population structure, in which agents choose their interactions individually, according to their (limited) ability to identify who they are interacting with. In this case, isolating the parasite is more difficult to achieve, since agents need to decide their interactions based on local information only. The structure of a population, as well as the code distribution of the agents, both affect the individual capacity to identify other agents. We believe this property to be crucial for the survival of populations, and a key issue for the development of an immune response. Acknowledgements This research was supported by “H2020-641321 socSMCs FET Proactive project”. The views expressed in this paper are those of the authors, and not necessarily those of the consortium.

References [1] Csete, Marie & Doyle, John (2011). Bow ties, metabolism and disease. *Trends in Biotechnology*, 22(9), 446-450.

## Dual Coding of Memory in Classical Texts

Han-liang Chang

University Chair Professor, Fudan University, Shanghai, China  
Professor Emeritus, National Taiwan University, Taipei, Taiwan

This paper discusses the theory of dual-coding and its application to information processing in memory by rereading a few classical texts which deal, at least partially, with the literary function of memory, such as Plato's *Timaeus* and *Symposium* and Aristotle's *On Memory and Recollection*. Remote as they may seem, these texts demonstrate how information in memory can be dually encoded, stored and retrieved as both visual and verbal representations, but once when language's meta-semiotic function is activated, all representation remains linguistic and thus coding's duality is compromised.

The Dual Coding Theory or Dual Code Theory (DCT), launched by Allan Paivio (1969, 1986, 2007), is a theory of memory which suggests that visual and verbal information act as two distinctive systems and humans are capable of storing information in either visual codes and/or verbal codes. The theory has been so popular in cognitive sciences that its usage has been stretched from its original formulation and its application extended to other fields, such as life science. For instance, Jesper Hoffmeyer and Claus Emmeche (1991) have proposed the notion of "code-duality" to account for life - a notion that consists of two different linguistic "codes", one digital and the other analog. Compared with Paivio's eclectic model, Hoffmeyer and Emmeche have perhaps rightly banished the dubious visual coding and reinstated the primary, if solitary, function of verbal language in representation. Since its inception nearly four decades ago, Paivio's assumption has been that "images and verbal processes are assumed to function as alternative coding systems, or modes of symbolic representation" (1969: 243), and the two interconnected processes perform independent functions in cooperative activity (2007). The verbal and image systems are correlated, as one can think of the mental image of an automobile and then describe it in words, or read words and then form a mental image. In brief, the DCT identifies three types of processing: (1) representational, the direct activation of verbal or non-verbal representations, (2) referential, the activation of the verbal system by the nonverbal system or vice versa, and (3) associative processing, the activation of representations within the same verbal or nonverbal system. More recently, Paivio has refined the processing to include five aspects: (1) verbal and nonverbal symbolic systems that cut across sensorimotor systems; (2) the representational units of each system; (3) connections and activation processes within and between systems; (4) organizational and transformational processes; and (5) conscious versus unconscious processes (2007: 33).

Dual-coding in memory is quite common in Plato and Aristotle. Suffice it to give two examples, one from each philosopher. In the prologue of the *Timaeus*, the speaker Critias comments on his ability of retelling a childhood story he had overheard to this effect: "Marvellous, indeed, is the way in which the lessons of one's childhood 'grip the mind,' as the saying is. For myself, I know not whether I could recall to mind all that I heard yesterday; but as to the account I heard such a great time ago, I should be immensely surprised if a single detail of it has escaped me.... [T]he story is stamped firmly on my mind like the encaustic designs of an indelible painting." (26C). The idea that memory trace appears as imagery is best illustrated by Aristotle in his *On Memory and Recollection*. Where the faculty of recollecting is an ability, a function of the soul, the so-called φαντάσια (phantasia or imagination), what is recollected are φαντάσματός (phantasmatos) (450a). Aristotle evokes two figures to "represent" what appears in our recollection: τύπος (typos) a print type, and γραφή (graphē) which means "writing" (450b16), both being doubly coded, first by "imagined" visual signs and then by linguistic signs.

By discussing the phenomenon of dual-coding in the classical texts, the paper attempts to bring into rapport the current discipline of Memory Studies with its ancient forgotten predecessors.

## Code Biology and the future of AI

Joachim De Beule  
Data Scientist at Engagor, R&D department, Belgium  
[joachim@engagor.com](mailto:joachim@engagor.com)

Artificial intelligence and cognitive science can be considered part of code biology at least to the extent that they aim to unravel the secrets of human intelligence and that this will require an understanding of the formation and workings of the neural codes. With the advance of big data technologies such as deep learning, several distinguished scientists and entrepreneurs have recently expressed their concern about the possibility of a super-AI surpassing human intelligence and taking over control. At one level, I wish to temper these expectations and refocus the debate by arguing that the new technology in fact does not provide solutions to some fundamental problems such as the ‘symbol-grounding problem’, which is directly related to the problem of the neural codes.

One of the main factors that made recent advances possible is the availability of huge amounts of human-annotated training data, mainly due to the success of social media. For instance, each time someone tags a friend in a picture posted on Facebook, data is produced that is used to train face-recognition algorithms. As a result, the performance of these algorithms today matches (and sometimes even surpasses) that of humans. But this would not have been possible without the continuous stream of human-annotated training examples, which is why I argue that the intelligence and codes being captured by the big-data crunching algorithms still essentially are *human* intelligence and *human* codes. At the same time I wish to point to the possibility that a novel intelligence is indeed emerging. The novel intelligence is not situated at the level of neural codes however, but at that of *cultural* codes.

In this contribution, I will discuss this possibility and analyze recent disruptive advances from the perspective of code biology, specifically from my earlier ‘sketch for a theory of evolution based on coding’.

In short, social media as well as many other emerging Internet technologies today allow individual people to extend into the social sphere and coordinate at a global level. This induces conventionalization dynamics that can lead to the formation of novel cultural codes. Although recent advances in AI do not in themselves lead to a novel, independent ‘super -intelligence’, they do accelerate this process and ultimately may lead to a major or meta-system transition giving birth to a genuinely novel intelligence or code-maker.



## Discussion on Code Hierarchies and Coarse Graining

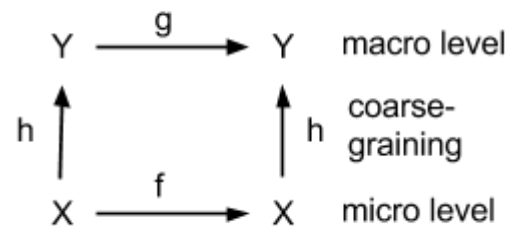
Peter Dittrich  
Bio Systems Analysis Group, Institute of Computer Science,  
Friedrich-Schiller-Universität Jena, Germany  
[dittrich@minet.uni-jena.de](mailto:dittrich@minet.uni-jena.de)

Joint discussion on: How do (organic) codes makeup hierarchies in living systems? How can we precisely describe and investigate this hierarchy? How can we use this for coarse graining the dynamics of living systems?

In the living world codes can be found at various scales, ranging from the molecular level over neuronal systems to the social level. These codes are not only operating at different spatial but also at different temporal scales. In this session we will discuss altogether whether the code view provides a general approach for coarse graining the complex dynamics of living systems. The genetic code indicates that this is at least partially possible, since the rules of the genetic code can be interpreted as a coarse-graining of a complex dynamical molecular system involving billions of molecular species.

Coarse graining from a mathematical point of view:

Given a dynamical process by a state transition function  $f: X \rightarrow X$  at a microscopic level with microscopic state space  $X$ , i.e.,  $x(t+1) = f(x(t))$ , a **coarse graining** is associated with two mappings: one mapping  $h: X \rightarrow Y$  that maps microscopic states  $X$  to macroscopic states  $Y$ , and one mapping  $g: Y \rightarrow Y$  that describes the dynamics on the macroscopic level, such that everything is consistent, i.e., the mappings commute as  $h(f(x)) = g(h(x))$  for all microscopic states  $x$  from  $X$ . Note that  $Y$  is usually much smaller than  $X$ .



How could this be related to a code view?

The **coarse-graining function  $h$**  describes which elements of the system (at the micro-state) are considered to represent which sign, symbol, meaning, or code-element. For example, for the genetic code,  $h$  maps (among others) chemical molecular species to words from  $\{A, C, G, T\}^*$ .

The **dynamics  $g$  on the macro-level** includes the code rules, i.e. the mapping between the “independent worlds”, which are both inside  $Y$ . For the genetic code,  $g$  includes the coding rules mapping triplets to amino acids.

Procedure:

The plan is to present a small set of guiding questions like the following at the beginning of the conference on Wednesday and ask for short contributions / answers to one or more of them to be presented at the discussion session on Friday:

1. How do (organic) codes makeup hierarchies in living system
1. How can we precisely describe and investigate this hierarchy?
2. How can we use this for coarse graining the dynamics of living systems?

## **Acoustic codes and bird dawn chorus**

Almo Farina  
Department of Basic Sciences and Foundations  
The University of Urbino, Campus Scientifico "Enrico Mattei"  
61029 Urbino - Italy  
[almo.farina@uniurb.it](mailto:almo.farina@uniurb.it)

Dawn chorus is one of the most spectacular behaviors of birds, and one about which there has been much speculation and the proposal of hypotheses without any definitive, univocal explanations. At first sight the acoustic patterns that emerge from a dawn chorus seems a chaotic blend of contemporarily songs from different species. When analyzed in details interesting patterns emerge that are the results of mechanisms operating at different temporal and spatial scale under the influence of environmental proxies like the composition of bird aggregations, the availability of resources and the vegetation structure of habitats.

A research conducted in five Mediterranean habitats during spring 2013 has found acoustic patterns at the level of community that for their regularity could be considered like ecological codes.

Dawn choruses start approximately one hour before the sunrise and stop suddenly at the sunrise for few minutes, then a further song activity starts again and decreases after a couple of hours.

Habitats rich in resources seem to have less differences between pre-sunrise and post-sunrise song activity and some vegetation patterns (like evenness) seem to have a relevant influence.

## Codon Usage in Circular and Comma-Free Codes

Elena Fimmel  
Mannheim University of Applied Science  
[e.fimmel@hs-mannheim.de](mailto:e.fimmel@hs-mannheim.de)

Lutz Strüngmann  
Mannheim University of Applied Science  
[l.struengmann@hs-mannheim.de](mailto:l.struengmann@hs-mannheim.de)

In 1957, Francis Crick suggested an ingenious solution for solving simultaneously the problem of amino acid coding and frame maintenance. The idea was based on the notion of comma-free codes. Even if the hypothesis of Crick revealed wrong, in 1996 Arquès and Michel discovered the existence of a weaker version of such codes in eukaryote and prokaryote genomes: the so called circular codes. Since then, circular code theory has provoked great interest and underwent a rapid development.

In this talk the codon usage in maximal comma-free or self-complementary circular codes will be discussed, i.e. we will investigate in how many of such codes a given codon participates. As the main (and surprising) result it will be shown that the codons can be separated in very few classes (three respectively six) with respect to their usage.

## Are coding sequences optimized?

Simone Giannerini  
University of Bologna, Department of Statistical Sciences  
[simone.giannerini@unibo.it](mailto:simone.giannerini@unibo.it)

*joint work with*  
Diego L. Gonzalez  
CNR/IMM Bologna Section  
and  
Greta Goracci  
University of Bologna, Department of Statistical Sciences

In this work we investigate the existence of universal optimizations in coding sequences. Usually, these take the form of correlations between nucleotides that are observed in many organisms and that might be related to error correction and energy optimization. We address the problem motivated by a mathematical model of the genetic code introduced in Gonzalez(2008) and further studied in Gonzalez et al.(2008, 2009), Giannerini et al.(2012). This new paradigm leads to the definition of dichotomic classes that can be seen as nonlinear functions of the information contained in a dinucleotide. Such classes, besides emerging naturally from the mathematical model, represent precise biochemical interactions. We use the dichotomic classes as a binary coding scheme for DNA sequences and study their mutual dependence by using suitable resampling techniques. We find universal strong short-range correlations between certain combinations of dichotomic classes. Remarkably, in some instances, we find that the sequences of dichotomic classes are less correlated than purely random sequences and this suggests that coding sequences can be seen as a minimum (or maximum) of some functional of the energy. The results indicate that the paradigms of information/communication theory are essential for the understanding of the organization of genetic information.

### References

- D. L. Gonzalez. The mathematical structure of the genetic code. In M. Barbieri and J. Hoffmeyer, editors, *The Codes of Life: The Rules of Macroevolution*, volume 1 of *Biosemitotics*, chapter 8, pages 111-152. Springer Netherlands, 2008.
- D. L. Gonzalez, S. Giannerini, and R. Rosa. Strong short-range correlations and dichotomic codon classes in coding DNA sequences. *Physical Review E*, 78(5):051918, 2008.
- D. L. Gonzalez, S. Giannerini, and R. Rosa. The mathematical structure of the genetic code: a tool for inquiring on the origin of life. *Statistica*, LXIX(3-4):143-157, 2009.
- S. Giannerini, D. L. Gonzalez, and R. Rosa. DNA, frame synchronization and dichotomic classes: a quasicrystal framework. *Philosophical Transactions of the Royal Society, Series A*, Vol. 370, Number 1969, 2987-3006, 2012.

## The Role of Code Systems in Linguistic Communication

Louis J. Goldberg  
State University of New York at Buffalo  
[goldberg@buffalo.edu](mailto:goldberg@buffalo.edu)

Inter-human communication enabled by both oral and written language is considered to be coded. Nevertheless, the various disciplines engaged in the study of linguistic communication do not fully comprehend the nature of biological coding systems and do not use the tenets of code biology as a framework for viewing the results of their studies or for developing testable models. This presentation is a beginning attempt to view language through the lens of code biology.

The basic code-elements of human oral linguistic communication are phonemes—the vowels and consonants of a language. Phonemes are articulated by the speaker and conveyed as ‘digital’ packets of acoustic waves that traverse the atmospheric space between humans. Phonemes in their manifestation as acoustic packets play the crucial role of conveying the acoustic signal from one human to another, but they are only one of several phonemic code-elements. For example, the problem in a hear-and-repeat paradigm of linguistic communication is to transform an acoustic phonemic code-element articulated by a sender (e.g. any one of the consonants) into the same consonant voluntarily repeated by a receiver. In essence, the consonant articulated by the sender enters the environment, goes through the brain of the receiver, is spoken by the receiver and comes back into the environment. The receiver-detected acoustic pattern has a particular structural signature. It is obvious that the medium of the body and brain cannot support travel of an acoustic signal. The information encoded in the pattern of the acoustic signal undergoes many transformations as it travels through the receiver’s auditory receptor system, to the cerebral cortex and then to the vocal musculature. Remarkably, the information contained in the pattern of the acoustic signal is conserved in the code-element transfer process through body and brain. The brain has transformed itself, along with its ancillary sensory receptor and musculoskeletal systems, into a code generating and receiving entity to enhance the effectiveness of inter-human communication.

The invention of writing introduced an additional set of code-elements into linguistic communication. These are the graphemes, the orthographic representations of phonemes. The manner in which graphemes as code-elements are transferred throughout the body and brain are similar to that described above for phonemes. Writing brought literacy into human social systems. Literacy is considered to be a uniquely human cross-modal cognitive process. Auditory phonological representations in brain (the phonemes of speech) become associated in brain with orthographic representations (the graphemes of reading and writing). To use the vague conceptual terminology of cognitive science and neurobiology, this requires the association of the processing stream of phonological information with the visual-object processing stream.

The exceptional cross-modal abilities of human language are currently of high interest to cognitive scientists. Consider that literate humans can hear-to-write (take dictation); read-to-write (copy); read-to-speak (give speech from text). Consider also that humans can accomplish these operations without external input, e.g., think-to-write, think-to-speak and think-to-think (self-to-self conversations). The code biology perspective leads to the proposition that these kinds of code-based, cross-modality operations are accomplished through the actions of an adaptor system in brain association cortex that governs the obvious rule-based coded correspondences among writing, reading and speaking. Although the words, code, encode and decode, are casually used in many publications concerning linguistic communication, an understanding of the coding processes as elaborated by code biology is not in evidence. In my view, the cognitive science and neurobiological approach to linguistic communication would be greatly enriched by an appreciation of the principles of code biology.

## On the origin of degeneracy in the genetic code

Diego L. Gonzalez  
[gonzalez@bo.imm.cnr.it](mailto:gonzalez@bo.imm.cnr.it)

The genetic code is the first molecular code identified in living organisms. Its very existence poses many fundamental questions about the origin and evolution of life. Here we present an analysis from a new perspective on this challenging problem. For this analysis we rely on both, a biological hypothesis for the origin of degeneracy, and a mathematical model describing exactly the degeneracy distribution of different variants of the genetic code. The biological hypothesis is based on primeval symmetry properties of nucleic acids and of molecular adaptors [1]. The mathematical model, instead, is based on redundant integer number representations systems, in particular, on binary, non-power, positional representation systems [2]. First, we analyse the degeneracy structure of the vertebrate mitochondrial genetic code [1]. Because this is the simplest variant among all known versions of the genetic code, different authors proposed it as a paradigmatic model for the early genetic code, that is, the code that preceded the universal genetic code characteristic of the Last Common Universal Ancestor (LUCA). Our tessera code, based on the above mentioned symmetry properties, is the natural candidate as the ancestor of the early code, that is, it is a parsimonious solution for the pre-early code. Second, we describe different steps that predated the appearance of the pre-early code until the so-called archetypical code [3,4], and beyond, near the origin of amino acid coding. Such steps strongly suggest that, as in many different physical examples, the evolution of the genetic code proceeded through successive symmetry breaks. These correspond to a progressive ambiguity reduction, that is, progressive decreasing degeneracy in amino acid coding. Finally, going forward in time, we can interpret within the same framework different schemes associated to post-transcriptional modifications in modern codes that putatively appeared at least after the establishment of the early code. From the viewpoint of code theory, we propose that the present forms of the genetic code are profoundly determined by physico-chemical constraints, mainly the symmetry properties of the molecules involved in primeval protein synthesis. Moreover, we observe a strong conservation of the degeneracy distribution that can be followed from the pre-early code to the present vertebrate mitochondrial genetic code and include the preservation of other relevant symmetry properties. Because the proposed solution for the pre-early code exhibit intrinsic features of error detection/correction systems, we hypothesize that the evolutionary pressure for decoding efficiency represents the main cause for its selection. However, we cannot exclude that other important biological functions might be associated to the conservation of the degeneracy properties of the genetic code.

### References

- [1] Gonzalez Diego L., Giannerini S., and Rosa R., "On the origin of the mitochondrial genetic code: Towards a unified mathematical framework for the management of genetic information", *Nature Precedings*, <http://dx.doi.org/10.1038/npre.2012.7136.1>, (2012).
- [2] Gonzalez Diego L., Giannerini S., and Rosa R., "On the origin of degeneracy in the genetic code", *preprint* (2015).
- [3] Gonzalez Diego L., "The Mathematical Structure of the Genetic Code", *in: The Codes of Life (book chapter 6)*, Marcello Barbieri Editor, Springer Verlag (2008).
- [4] Watanabe, K. and Yokobori, S., "How the early genetic code was established? -inference from the analysis of extant animal mitochondrial decoding systems-", in *Chemical Biology of Nucleic Acids: Fundamentals and Clinical Applications*, RNA technologies, Erdmann, V., Markiewicz, W. and Barciszewski, J., eds., chap. 2, 25-40 Springer Verlag, (2014).

## The relational basis of molecular codes - Towards hierarchies of molecular codes

Dennis Görlich  
Institute of Biostatistics and Clinical Research,  
Westfälische Wilhelms-Universität Münster, Germany  
[Dennis.goerlich@ukmuenster.de](mailto:Dennis.goerlich@ukmuenster.de)

The existence of codes in biology, such as the genetic code, is a fact that enables us to search for other codes in biological systems. Some other cellular subsystems have been proposed to implement an encoded relationship between molecular “worlds”, e.g. the histone code. We may study and identify each code individually, i.e. without its relation to other codes, but these relations may exist and can be described formally. Recently, I proposed the idea of the relational basis of molecular codes (Görlich, 2014) which formalizes the relations between molecular codes mathematically based on the formalization presented in (Görlich & Dittrich, 2013). Two types of code relations can be identified (i) a linkage of codes, illustrating the situation that codes are executed subsequently as part of an information processing cascade (ii) a nesting of codes. Both types enhance a systems code capacity, i.e. the number of mappings that could be realized. Code linkage can be present in two subtypes (a) where the meanings of a first codes, can be used as signs in a second code, and (b) a setting where the meanings of a first code have an effect on the context of the second code, i.e. the first code controls the coding rules of the second code. I will discuss these situations, sketch proofs for the mathematical properties of the relations, and propose the concept of hierarchies of codes, or code networks.

Görlich D (2014) The relational basis of molecular codes. *Biosemitotics* 7:249-257.

Görlich D, Dittrich P (2013) Molecular Codes in Biological and Chemical Reaction Networks. *PLoS ONE* 8(1): e54694.

## BDA-generated Models of the Genetic Code in current and ancient genetic codes

M. Gumbel<sup>1</sup>, E. Fimmel<sup>1</sup>, A. Danielli<sup>2</sup>, L. Strüingmann<sup>1</sup>

<sup>1</sup>M. Gumbel ([m.gumbel@hs-mannheim.de](mailto:m.gumbel@hs-mannheim.de)), E. Fimmel ([e.fimmel@hs-mannheim.de](mailto:e.fimmel@hs-mannheim.de)),  
L. Strüingmann ([l.struengmann@hs-mannheim.de](mailto:l.struengmann@hs-mannheim.de))

Faculty of Computer Science, Mannheim University of Applied Sciences,  
Paul-Wittsack-Str. 10, D-68163 Mannheim, Germany

<sup>2</sup>A. Danielli ([alberto.danielli@unibo.it](mailto:alberto.danielli@unibo.it))

University of Bologna, Department of Pharmacy and BioTechnology, Via Irnerio 42, 40126 Bologna, Italy

In this talk we introduce the concept of models of the genetic code which are based on binary dichotomic algorithms (BDAs) [1][2]. A BDA divided the set of 64 codons into two disjoint classes of size 32 each and provides a generalization of known partitions like the Rumer dichotomy [3] or the Parity class dichotomy [4]. We analyze what partitions can be generated when a set of different BDAs is applied sequentially to the set of codons. An algorithm is developed that scans for interesting BDA-generated models. The search revealed that those models are able to generate code tables with very different number of classes ranging from 2 to 64. It is analyzed whether there are models that map the codons to their amino acids. A perfect matching is not possible. However, there are models that describe the standard genetic code with only little errors. There are also models that map all 64 codons uniquely to 64 classes showing that BDAs can be used to identify codons precisely. This could serve as a basis for further mathematical analysis using e. g. coding theory. The hypothesis that BDAs might play a role at the decoding center of the ribosome is discussed and first results on BDAs in the context of ancient genetic codes are presented.

### REFERENCES

1. M. Gumbel, E. Fimmel, A. Danielli, L. Strüingmann: On Models of the Genetic Code generated by Binary Dichotomic Algorithms. *Biosystems* 128 (2015) 9-18.
2. E. Fimmel, A. Danielli, L. Strüingmann: On dichotomic classes and bijections of the genetic code, *J. Theor. Biology*, Volume 336, Pages 221-230 (November 2013).
3. Y.B. Rumer: Systematization of codons in the genetic code, *Doklady Akademii nauk SSSR* 1969 Aug; 187(4):937-8.
4. D.L. Gonzalez: The Mathematical Structure of the Genetic Code, In Barbieri (Ed.): *Codes of Life: The Rules of Microevolution*, Springer, 2008



## **Synthesis of living systems : perspectives from artificial life**

Martin M Hanczyc  
University of Trento  
Centre for Integrative Biology (*CIBIO*)  
Trento, Italy  
[Martin.Hanczyc@unitn.it](mailto:Martin.Hanczyc@unitn.it)

My work is focused on understanding the fundamental principles of living and evolving systems through experimental science. To this end I build synthetic systems where dynamic life-like properties emerge when self-assembled systems are pushed away from equilibrium. I will present an experimental model of bottom-up synthetic biology: chemically-active oil droplets. This system has the ability to sense, metabolize and the potential to evolve. Specifically, I will present how sensory-motor coupling can produce chemotactic motile droplets. In addition I will present how we add information to our supramolecular structures in the form of ssDNA that then governs the assembly and disassembly of higher order architectures.

## On the logical necessity for a manufacturing code in self-fabricating systems

Jan-Hendrik Hofmeyr  
Centre for Studies in Complexity, University of Stellenbosch,  
Stellenbosch 7600 South Africa  
and  
Wissenschaftskolleg zu Berlin, Berlin 14193, Germany  
[jhsh@sun.ac.za](mailto:jhsh@sun.ac.za)

We, the living, all do something astonishing that no non-living thing can do. Unlike, say, a car, which needs you or a mechanic to replace or repair its parts, we can fabricate all our parts by ourselves. Self-fabrication underlies all those properties so often used to define life: growth, metabolism, reproduction, development, adaptation, even evolution. Without explaining self-fabrication we can explain neither life itself, nor its origin. In traditional fabrication, as carried out by humans, there is an external agent, the artisan. In Aristotelian terms, such an agent is the efficient cause. In self-fabrication, however, there is no external agent; all the efficient causes are produced within the system. Such a system is, therefore, closed to efficient causation. But such closure requires an efficient cause to be produced by another efficient cause, which in turn must be produced by another efficient cause, which seems to lead to an infinitely regressing chain of efficient causes. I shall present a formal model for self-fabrication that resolves this conundrum. It models the fundamental and universal biochemical process of polymerization as the creation of symbol strings in a formal language by production rules, which are themselves created as symbol strings from within the language in order to ensure closure to efficient causation. But, just as a linear polymer must fold into a specific three-dimensional structure before it can become functional, so a symbol string that describes a production rule must in some way acquire semantic meaning before it can do its job. This implies that just as the correct folding of a polymer requires the appropriate chemical context, so a production rule string can only acquire meaning in an appropriate linguistic context. Through a series of logical steps I build on this formal scaffolding to construct a complete model of a self-fabricating formal language. I show that avoiding the infinite regress of rules that produce rules that produce rules, and so on, necessitates the separation of the description of the sequence of symbols from the construction of the sequence; this in turn necessitates the prior encoding of the sequence and the subsequent translation of the encoded form of the sequence by a specialised set of rules that are also produced within the system. Such a coding system has been called a manufacturing code. The beauty of this linguistic model of self-fabrication is not only that it maps perfectly onto cell biochemistry as we know it, but that it demonstrates that the problem of closure to efficient causation in linear polymer systems, the key to self-fabrication of all living things, can logically be solved in only one way.

## Encoding Space

Tim Ireland

Senior Lecturer, Programme Leader MA Architectural Design

Leicester School of Architecture

De Montfort University Leicester, UK

[tireland@dmu.ac.uk](mailto:tireland@dmu.ac.uk)

Space like many concepts (such as those of information and mind) escapes concrete definition. Taking a biological stance on the matter space is perceived a dynamic emergent manifold which, like a dangling carrot hanging before us but beyond our reach, evades comprehension. As such, trying to define what space is may be deemed to be a waste of ... space. However, space is a fundamental property of life (Hall 1966). Biotic beings are spatial and exist in a predominantly spatial manner, affecting and affected by those things they share the world with; and although one being may be deemed of no significance to another it holds potential significance. The spider's web, for example, works because flies cannot detect its presence. The web exists even though the fly doesn't perceive it; until, that is, it is too late for the fly.

As a relationship which is produced supra-subjectively across agents, space has objective properties. However it is not an object per se, but a 'form', as Kant claimed of 'sense perception' because it is tangible in the sense that it can be perceived - and can be acted upon as a sign. At this most primal level it is a pattern of interaction. As an artefact, formed through an organism's capacity to affect and manipulate its environment, space is objectified. The definitive manifestation of physical space is thus an artefact, which embodies the spatiality of the organism that created it. Having progressed from congregating around fire, humankind constructs buildings serving purposes beyond basic physiological needs, such as cultural and personal expression. These patterns are established through codes-of-being, propagated in a dialectical process of interaction between organism and environment, effecting artefacts such as nests, structures and buildings.

The French Marxist philosopher and sociologist Henri Lefebvre recognised the mental and physical aspects of space as intertwined and mediated through what he called spatial-practice; referring to habitual tendencies cast into the artefacts and structures we inhabit. In other words, lived-space is the assimilation of physical and mental space. His spatial-code expresses a tri-dialectic process whereby space is created and creative effected through the relations between subjects, their space and surroundings, meaning space is social because it unfolds through interaction. As a social product, he states, "[s]pace is neither a 'subject' nor an 'object' but rather a social reality - that is to say, a set of relations and forms" (Lefebvre 1995: 116). Emphasising the social dimension of being in the world Lefebvre stresses interaction is both mental and physical. "Space is social morphology: it is to lived experience what form itself is to the living organism, and just as intimately bound up with function and structure" (Lefebvre 1995: 94). His model provides a basis for scrutinising spatial practice but doesn't provide a means for, as the genetic code does, generation, because it does not define rules which may be interpreted and built upon. So if space is a product of mediating mental and physical aspects of being what is the underlying code that enables this mapping? Is there an underlying elemental code on which codes-of-being are founded? If so would it facilitate designing, and thus enhance the built environment, and how might this capacity be applied?

## Genetic Code Tool (GenTool)

Elena Fimmel, Lutz Strüngmann and  
Kristian Kraljic  
Institute of Applied Mathematics  
Faculty for Computer Sciences Mannheim  
University of Applied Sciences  
68163 Mannheim, Germany  
[k.kraljic@kooperationen.hs-mannheim.de](mailto:k.kraljic@kooperationen.hs-mannheim.de)

The Genetic Code Tool (GenTool) is a tailor-made software tool, to support and simplify the creation and mathematical analysis of sets and sequences of genetic code tuples, such as trinucleotide codes and their corresponding amino acids. The main purpose of the tool is to provide a simple and modern user interface and a tool-aided approach, to create, transform and visualize sequences or sets of tuples, as well as to perform tests and analysis on those. Written in Java the tool is easy to enhance via an open interface, to add an arbitrary number of custom transformations, splits, tests and analyses, as well as new controls to input and visualize genetic code sequences and sets.

## Catch and release: attachment and dispersal codes in a bacterial biofilm

Stefan Kühn and Ruenda Loots  
Dept. of Biochemistry, University of Stellenbosch,  
Stellenbosch, 7600, South Africa  
[stefankuhn21@gmail.com](mailto:stefankuhn21@gmail.com)

Biofilms are naturally occurring, surface-bound aggregates of microbes. Bacteria, Archea, fungi, algae, and protozoans are able to accumulate within an autogenic three-dimensional matrix of extracellular polymeric substances. Such aggregates have been independently called 'microbial cities', 'secret societies', and even the 'bacterial twitter'. The need for communication both inter- and intra-species is paramount if such densely populated, highly organized 'societies' are to survive and thrive. However, before our humble microbes blossom into an entire city, the first brick must be laid. This necessitates a mechanism by which a microbe may sense its environment, relay the signal, and translate it into action. We propose that this mechanism, despite experimental difficulties, exists and can be explained from the perspective of code biology. The system wherein the microbe - *P. aeruginosa* - senses its environment and translates that signal, via an intricate array of biomolecules, to states of attachment, growth, or dispersal are the attachment and dispersal codes. To aid us in the explanation of these codes we draw upon the body of knowledge provided by Barbieri and his work on organic codes. Since the term 'code' has seen a number of uses - not all of which conform to the precepts of an organic code as provided by code biology - we shall endeavour to test the precepts of the bacterial communication code against those established by code biology. Drawing on our current knowledge of *P. aeruginosa* biofilms we will narrate the attachment, development, and dispersal of just such a biofilm from within the context of an organic code. Further, we will show that the environment, signalling molecule, and effect threesome of such bacterial attachment and dispersal systems conform to the sign-adapter-meaning trichotomy as laid out in code biology. A particular change in the microbe's environment is sensed and acts as an organic sign which is then relayed by a specialized signaling molecule that acts as an adapter. This adapter molecule then translates the sign into attachment or dispersal - the biological meaning.

## The Coding Coenzyme Handle Hypothesis revisited

Ádám Kun  
[kunadam@elte.hu](mailto:kunadam@elte.hu)

According to the Coding Coenzyme Handle Hypothesis, the genetic code first served to reliably attach amino-acids to ribozymes, so the amino-acids can offer their chemical versatility to the ribozymes. The amino-acids are attached to small oligonucleotides (handles), which in turn can attach to ribozymes via hydrogen-bonds. Complementary triplets forming kissing-hairpins can offer a strong enough binding. If amino-acid were coenzymes fostering catalysis, then (1) there are not many sites on a ribozyme onto which the handles could attach, and (2) we find mostly codons of amino-acids with known catalytic importance at these sites. We investigate these questions by analysing the secondary structures of random RNAmolecules and counting the number of attachment sites as well as their identity. It turns out that the codons for lysine, asparagine, glutamine, isoleucine, leucine and phenylalanine are frequently found in single-stranded regions of RNSs, thus offering potential attachment sites to the coenzyme-handles. According to some theories, only the middle nucleotide or the first two nucleotides were used in coding in the past. If we accept that only the middle nucleotide is important, then codons with A and U in the middle are the ones that could be good handles. We find many of the catalytically important amino-acids (histidine, aspartic acid) in the A column. This is good for the theory! However, we find apolar amino-acids in the U column. There is some debate in the literature about the early role of amino-acids, and catalysis is only one of them. They could have been employed in a structural role, helping ribozyme to fold properly, as well as helping membrane functions (RNA cannot submerge into membranes). Our results suggest that according to the extant code coupled with the Coding Coenzyme Handle Hypothesis, they could serve both roles.

## Computational modelling of the evolution of symbolization

Barry McMullin  
Dublin City University, Ireland  
[barry.mcmullin@dcu.ie](mailto:barry.mcmullin@dcu.ie)

The paradigmatic example of a biological code is the relation between the nucleotide sequence in mRNA and amino acid sequence in a corresponding protein. It is of the essence of this code that its physical realisation is reflexive: the key protein enzymes responsible for the specificity of the code (the t-aminoacylsynthetases) must themselves be self-consistently coded in corresponding genes in order for the coding system to be sustained (and ultimately, to be reproduced). I will present a schematic, but fully functional, computer model (or, more precisely, “virtual world”) which attempts to capture this key characteristic of an arbitrary, but self-referentially bootstrapped and sustained, coding system. This is based on realisation of the generic architecture for machine self-reproduction first formulated and articulated by John von Neumann in the late 1940’s. While the model is simple, it is still conceptually powerful, and allows practical investigation of real evolutionary phenomena perturbing the initially defined coding system in complex ways. Some concrete examples of these phenomena will be presented.

## **Futile metabolic cycles and memory generation**

Chris Ottolenghi  
Service de Biochimie Métabolomique et Protéomique  
Hôpital Necker Enfants Malades  
149 rue de Sèvres, 75015 Paris, France  
[chris.ottolenghi@parisdescartes.fr](mailto:chris.ottolenghi@parisdescartes.fr)

In individual organisms, metabolic cycles mediate a range of functions such as substrate transport across biological boundaries (e.g., the malate-aspartate shuttle); energy storage (e.g., creatine phosphate); heat production (e.g., mitochondrial uncoupling proteins); or complex reactions involving ‘catalytic’ metabolites (e.g., acetyl-CoA combustion via the Krebs cycle). Metabolic cycles are termed “futile” when they do not show any net effect except (supposedly) wasting energy. Indeed, futile cycling has been implicated in several mechanisms of pathology. But what for the apparently futile cycles that occur under physiological conditions? I will review examples focusing on distinct lipogenesis-lipolysis cycles. Available data will be discussed in relation to the concepts of cell memory generation, stemness, and transdifferentiation.



## Collectives, Codes and Conventions

Daniel Polani  
Department of Computer Science  
University of Hertfordshire  
Hatfield AL10 9AB, United Kingdom  
[daniel.polani@gmail.com](mailto:daniel.polani@gmail.com)

We discuss the role of information in self-organization of collectives, information-limited acquisition of information about the world and how agents can improve the quality of their (limited and incomplete) world description by "agreeing" on conventions to describe it. It turns out that collectives indeed can capture more structural properties about the environment than the individual if information exchange is limited. In other words, information-limited codes may sharpen agents' perception of the structure of their environment.

## I/O Maps and Codes in Regulatory Systems

Sonja Prohaska  
Computational EvoDevo  
Institute of Computer Science  
University of Leipzig  
[sonja@bioinf.uni-leipzig.de](mailto:sonja@bioinf.uni-leipzig.de)

Regulation is inherently dynamic and always involved with the transmission of information from inputs to outputs. To keep inputs and outputs from mixing, they need to be of different types. Consequently, an mapping of values for the input variable onto corresponding output values can be formally represented as a typed I/O map. This map encapsulates a causal relation and may or may not constitute a code.

In this contribution I will discuss the potential relevance of the computer science concepts "indirection" and "arbitrariness" for biological codes.

Furthermore, I will elaborate on the importance of organizational or logical structures in the input values and the mapping. To complete the list of conceptual elements of regulatory systems, I shall mention "feedback" and "memory" and use the well studied regulatory system, the lac operon, as an example.

## Olfactory Coding - from Odor Molecules to Brain Activity

Silke Sachse  
Olfactory Coding Group  
Max Planck Institute for Chemical Ecology  
Department of Evolutionary Neuroethology  
Hans-Knoell-Strasse 8  
D-07745 Jena  
Germany  
[ssachse@ice.mpg.de](mailto:ssachse@ice.mpg.de)

We are investigating how odors are coded and processed in the *Drosophila* brain to lead to a specific odor perception. The basic layout of the first olfactory processing centers, the vertebrate olfactory bulb and the insect antennal lobe, is remarkably similar. Odors are encoded by activated glomeruli in a combinatorial manner. *Drosophila melanogaster* provides an attractive model organism for studying olfaction, as it allows genetic, molecular and physiological analyses. The talk will summarize our recent insights into coding strategies of ecologically relevant odors yielded by morphological and functional analysis of the neuronal populations present in the antennal lobe and aim to find the link to odor-guided behavior.

## Codes in RNA - How to design RNA Structures

Peter F. Stadler  
Bioinformatik, Inst.f.Informatik, Univ.Leipzig  
Haertelstrasse 16-18, D-04107 Leipzig, Germany  
[peter.stadler@bioinf.uni-leipzig.de](mailto:peter.stadler@bioinf.uni-leipzig.de)

The sequences of RNA molecules implicitly determines the molecules' structure. Largely, this encoding is determined by the thermodynamics of RNA folding, to a certain extent also by the kinetics of the folding process and by protein that interact with the RNA.

Meanwhile we understand the rules that link sequence and structure of RNAs sufficiently well to use them for the rational design of RNAs with intricate functions: riboswitches are abundant regulators of both transcription and translation in procaryotes that rely on environmentally triggered structural changes.

After an introduction in the simple rules of the RNA structure "code", the presentation will focus on the inverse folding problem and practical design strategies for functional RNAs.

## Implementing Molecular Logic Gates, Circuits, and Cascades Using DNAzymes

Darko Stefanovic  
[darko@tau.cs.unm.edu](mailto:darko@tau.cs.unm.edu)

The development of electronic digital logic was one of the greatest technological achievements of the 21st century, and exponential increases in the computational power of commercially-available microprocessors meant that electronic computers are now ubiquitous and indispensable in the modern world.

Contemporaneous advances in molecular biology made it clear that information processing is a fundamental capability of all biological systems. Subsequent rapid progress in that field progressed in parallel with the development of consumer electronics, and elucidated many of the mechanisms behind

biological information processing. Given that the information processing capabilities of biological systems were evolved over millions of years, it is fascinating to consider whether we can construct synthetic molecular computers that may be more compact and more robust than their natural or electronic counterparts. When we say that molecules compute, what we usually mean is that an assembly of molecules detects certain inputs, typically the presence or absence of other molecules, and responds by producing one or more output signals, which may take the form of the release of an output molecule or the generation of a detectable fluorescent signal. The goal of a molecular computer scientist is to engineer the intervening molecular system so that the pattern of output signals is related to the pattern of input signals by the desired logic function. From an unconventional computing perspective, the development of molecular computers offers intriguing possibilities to implement extremely low-power computation and to implement autonomous computational systems that can survive and thrive in environments hostile or inaccessible to silicon microprocessors, such as within the bloodstream or within living cells. The compact nature of DNA has been previously exploited to demonstrate high-density information storage, but in our context the fact that billions of molecules exist in each experimental system may make it feasible to execute massively parallel computations in a very small volume, or to implement novel computational architectures that compute using the dynamics of interactions between molecular circuit components.

Our experimental work focuses on catalytic nucleic acid chemistry, in particular, DNAzymes (deoxyribozymes), which are DNA-based enzymes that can cleave or combine other nucleic acid strands. DNAzymes are now known to occur in nature, and the known DNAzyme catalytic motifs have been isolated in *in vitro* evolution experiments. We turned DNAzymes into logic gates by augmenting them with up to three input-binding modules that regulated the catalytic activity of the DNAzyme based on the pattern of input strands observed in the solution. The cleavage reaction catalyzed by the DNAzyme served as the reporting channel, and we exploited the combinatorial chemistry of DNA to enable us to build systems that processed multiple signals simultaneously in a single solution, with the different information streams tagged by different DNA sequences. Thus, each DNAzyme unit implemented a logic gate with up to three inputs, and we constructed a set of such gates complete for Boolean logic. I will review our designs for DNAzyme-based molecular computers, their integration in large-scale parallel gate arrays exhibiting sophisticated logical and temporal behaviors, and our recent attempts to diversify into sequential logic cascades. Our early approach to molecular computing included the first reported complete set of nucleic acid-based logic gates, which we then used to produce autonomous molecular computing systems that implement well-known logic circuits such as adders and large-scale game-playing automata. I will then discuss current work in which we have extended this approach to achieve signal propagation in DNAzyme signaling cascades and have begun to apply these new techniques to biodetection applications.

## **Dinucleotide comma-free codes and their connection to dinucleotide circular codes**

Lutz Strümgmann  
Mannheim University of Applied Science  
[l.struengmann@hs-mannheim.de](mailto:l.struengmann@hs-mannheim.de)

Elena Fimmel  
Mannheim University of Applied Science  
[e.fimmel@hs-mannheim.de](mailto:e.fimmel@hs-mannheim.de)

The presence of circular codes in mRNA coding sequences is postulated to be involved in informational mechanisms aimed at detecting and maintaining the normal reading frame during protein synthesis.

Most of the recent research is focused on trinucleotide circular codes. However, also dinucleotide circular codes are important since dinucleotides are ubiquitous in genomes and associated to important biological functions. In this talk a group theoretic approach for investigating maximal dinucleotide comma-free codes will be shown and symmetry properties of such codes will be highlighted. Moreover, we describe a construction principle for such codes, their close connection to maximal dinucleotide circular codes and provide a graph representation that allows to visualise them geometrically. The results obtained can be interpreted as a weaker formulation of the well known hypothesis by Crick about frame-shift-detecting codes without commas.

## **Towards major evolutionary transitions theory 2.0**

Eörs Szathmáry

Center for the Conceptual Foundations of Science, Parmenides Foundation,  
Kirchplatz 1, D-82049, Munich, Germany;

MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group,  
1c Pázmány Péter, H-1117 Budapest, Hungary

[szathmary.eors@gmail.com](mailto:szathmary.eors@gmail.com)

The impressive body of work on the major evolutionary transitions in the last twenty years calls for a reconstruction of the theory, although a two-dimensional account (evolution of informational systems and transitions in individuality) remains. Significant advances include the concept of fraternal and egalitarian transitions (lower level units like and unlike, respectively). Multilevel selection first without, then with the collectives in focus is an important explanatory mechanism. Transitions are decomposed into phases of origin, maintenance and transformation (i.e. further evolution) of the higher-level units, that helps reduce the number of transitions in the revised list by two so that it is less top-heavy. After the transition units show strong cooperation and very limited realized conflict. The origins of cells, the emergence of the genetic code and translation, evolution of the eukaryotic cell, multicellularity, and the origin of human groups with language are reconsidered in some detail in the light of new data and considerations. Arguments are given why sex is not in the revised list as a separate transition. Some of the transitions can be recursive (e.g., plastids, multicellularity) or limited (that share the usual features of major transitions without a massive phylogenetic impact, such as the micro-and macronuclei in ciliates). During transitions new units of reproduction emerge, and establishment of such units requires high fidelity of reproduction (as opposed to mere replication).

## Umwelt codes exemplified by Umwelt alignment in corvids

Morten Tønnessen,  
University of Stavanger  
[mortentoennessen@gmail.com](mailto:mortentoennessen@gmail.com)

I have previously suggested that there are different *Umwelt codes*, which can be categorised as either CODE<sub>fix</sub> (fixed Umwelt codes) or CODE<sub>flex</sub> (flexible Umwelt codes). While neural codes are examples of CODE<sub>fix</sub>, ecological codes are instances of CODE<sub>flex</sub>. Some corvids, including crows, evidently prosper in part due to their relationships with human settlements and anthropogenic food sources. However, actual human-corvid relationships are typically somewhat distanced - likely because, for one thing, corvids are often treated by humans as pest species. In this paper I will look into the *Umwelt alignment* (cf. Tønnessen 2014) between corvids including crows on one hand and human beings on the other. My hypothesis is that Umwelt alignment must typically involve ecological codes (Umwelt codes).

What is Umwelt alignment? In dictionaries, “alignment” signifies an adjustment to a line, or arrangement into a straight line; a state of agreement or cooperation; the proper positioning or state of adjustment of parts in relation to each other; etc. Crucially, alignment can denote processes or states of fitting-in with others. If we define Umwelt alignment as the process of adjustment by one creature to the presence and manifestation of other Umwelt creatures (and further, to abiotic Umwelt objects and meaning factors), we realize that every Umwelt dweller on this planet conducts Umwelt alignment on a regular basis, as manifested over time in concrete functional cycles (see Uexküll 2010 [1934/1940]: 49).

Not all Umwelt alignment is mutual and cooperative. If Umwelt alignment is a universal phenomenon, then there must also be Umwelt alignment among natural enemies, and among competitors. *Keeping a certain distance* can be seen as emblematic of Umwelt alignment. The spatial distribution of specimens is central in human and animal social life as well as in terms of human ecology and general ecology, and is relevant in the current context to the extent that spatial distribution is arranged by way of Umwelt creatures with deliberate adjustment to the presence of others. Here we observe not only various forms of natural (autonomous) Umwelt alignment, but further instances of coerced Umwelt alignment, a phenomenon which is enforced and motivated more or less exclusively by human utility.

The regulatory mechanism of Umwelt alignment thus ranges from symbiotic strategies to more competitive forms of coexistence. In all cases, however, various forms of synchronicity are key. A study of Umwelt alignment between corvids and humans can expectedly shed light on human-corvid co-evolution, corvid Umwelten, and current human ecology.

### References

- Tønnessen, Morten 2014. Umwelt Trajectories. *Semiotica* 198 (Special Issue on zoosemiotics, guest-edited by Timo Maran): 159-180.
- Uexküll, Jakob von 2010 [1934/1940]. *A foray into the worlds of animals and humans - with A theory of meaning* (Posthumanities 12), Joseph D. O’Neil (trans.). Minneapolis & London: University of Minnesota Press.

**Acknowledgement:** This work has been supported by the research project “Animals in Changing Environments: Cultural Mediation and Semiotic Analysis” (EEA Norway Grants/Norway Financial Mechanism 2009-2014 under project contract no. EMP151).



## Information Coding in the Brain

Jochen Triesch  
Frankfurt Institute for Advanced Studies  
Ruth-Moufang-Straße 1,  
60438 Frankfurt am Main, Germany  
[triesch@fias.uni-frankfurt.de](mailto:triesch@fias.uni-frankfurt.de)

How does the brain encode information, i.e., what is the neural code? And how does the brain acquire the code(s) that it uses? Although these questions have always been a central topic of Neuroscience research, our current understanding is still very limited. I will first give an overview of the coding problem from the perspective of Neuroscience. Then I will focus on the efficient coding hypothesis, which claims that animal brains exploit redundancies in their sensory inputs to encode information more efficiently. Finally, I will discuss a recent extension called active efficient coding, which generalizes the theory to so-called active perception, which includes movements of the sense organs.