International Workshop on Computational Modelling

14 May, 2019, 16:00 to 20:00
Warsaw University of Technology, main building, Plac Politechniki 1, room 206

Contents

Turing, Ashby and "the Action of the Brain": Form and Function in Modelling the Mechanisms of Cognition

Hajo Greif

Not very much has been written to date on the relation between Alan M. Turing and W. Ross Ashby, besides citing and briefly discussing a letter from Turing to Ashby, in which Turing suggested using an early digital computer for "producing models of the action of the brain" (Turing 1946). Given the personal acquaintance between Turing and Ashby, and given the partial proximity of their research fields, an additional historical insight into the interactions between their views should be possible and worthwhile (for previous inquiries compare, for example, Asaro 2011; Dewhurst 2018). We suggest that a comparative view of Turing's and Ashby's work will also help to address a systematic question in the inquiries into human cognition: What is the relation between "producing models of the action of the brain" and the formal, symbol-based characteristics of the "instructions" on which possible computational models are based?

Reconstructing some of the most remarkable commonalities and differences between Turing's and Ashby's work, this inquiry will provide a perspective towards resolving the seemingly strict dichotomy between the formal-symbolic nature of modelling in Artificial Intelligence (AI) and the principle of embodiment contended by AI critics and proponents of "Nouvelle AI".

Both Turing and Ashby believed that "the action of the human brain" can be subject to a method of modelling that casts action in a strict mathematical description and breaks it down into simple routines that can be implemented in some kind of machine. However, they differed in terms of how and to what purpose that mechanical modelling is accomplished: First, the criterion of the model being "mechanical" was understood differently. The primary though not the exclusive focus was on formal versus material modelling respectively, which can be detected in their diverging interpretations of the role of mathematical methods in modelling. Second, Turing's and Ashby's interpretations of the models' target systems diverged: Ashby was concerned with adaptive behaviours of brains and other systems, their functions and their relationships to their environments (Ashby 1960). In doing so, he explicitly operated under a Darwinian paradigm. Turing's approach in (1952), in turn, built on the non-Darwinian account of morphogenesis developed by Sir D'Arcy Thompson (1942). His take on morphogenesis, along with his proto-connectionist ideas, were primarily concerned with the amenability of a variety of phenomena to computational modelling.

Either way, however, the primary target of modelling the "action of the brain" envisioned by the two authors might not refer to higher-order cognitive functions and symbolic representation, but to basic forms of cognitive organisation and adaptive functions of the brain as a biological organ, respectively. Hence, for entirely distinct reasons, Turing and Ashby appeared similarly indifferent to what counts as the key characteristic of human thought in the cognitive sciences and classical AI. By reference to contemporary approaches to cognition as being embodied and environmentally situated, it can be shown how this seeming indifference might prove instructive. Neither is there a need to draw a "fairly sharp line between the physical and the intellectual capacities of a man", as Turing believed, nor does a selectionist argument foreclose a finely grained and biologically defensible view of the action of the brain as a necessarily embodied, but not necessarily representational phenomenon.

References

Asaro, P. M. Computers as models of the mind: On simulations, brains, and the design of computers. In *The Search for a Theory of Cognition. Early Mechanisms and New Ideas*, S.
Ashby, W. R. *Design for a Brain. The Origin of Adaptive Behaviour*, 2 ed. John Wiley & Sons, New York/London, 1960.
Dewhurst, J. British cybernetics. In *The Routledge Handbook of the Computational Mind*, M. Sprevak and M. Colombo, Eds. Routledge, London, 2018, pp. 272–282.
Franchi and F. Bianchini, Eds. Rodopi, Amsterdam/New York, 2011, pp. 89–114.
Thompson, D. W. *On Growth and Form*, 2 ed. Cambridge University Press, Cambridge, 1942.
Turing, A. M. Letter to W. Ross Ashby of 19 November 1946 (approx.). *The W. Ross Ashby Digital Archive*.
Turing, A. M. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society, B* 237 (1952), 37–72.

Non-Turing models of computation as a hypothetical basis for modeling some mental activities

Paweł Stacewicz

1. I call a **non-Turing** model of computation (NTMC) any model that is not extensionally equivalent to the Universal Turing Machine model (Turing 1936). Each model of this kind allows to solve a broader class of problems than the Turing model (it is extensionally broader (Copeland 2002), (Ord 2006)).

2. The informatic motivation to extend the Turing model stems from the fact that in this model there are **uncomputable** problems, such as the halting problem or the problem of diophantic equations (Harel 1987). The unsolvability of such problems requires, first, an understanding (why they are not solvable) and, second, an indication of real or hypothetical methods of overcoming them (how they can be reformulated and/or solved).

3. One of the strategies for defining NTMC models, also called models of **hypercomputation**, consists in modifying at least one of the distinguishing features of the Turing model: a) *discreteness* (digitality), b) *finiteness* (finite number of operations performed in finite time), and c) *determinism* (a strictly defined data processing scheme).
Modification of one of the above mentioned features leads to the following models: a') *analogue-continuous*, b') *infinitistic*, c') *non-deterministic* models, respectively. Modification of more features leads to mixed models (Shannon 1941), (Shagrir 2004), (Ord 2006), (Burgin, Dodig-Crnkovic 2013).

4. In the field of computational models of **mental** activities (such as perception, reasoning or learning), the dominant systems are those compliant with the Turing model of computation, i.e. systems which, at the elementary level of mathematical description, are equivalent to a certain Turing machine (in practice: machine with a finite tape). These include both symbolic (e.g. rule-based systems) and connectionist systems (e.g. perceptron-type neural networks). All of them are implemented using programs for digital computers (sometimes controlling certain physical systems, e.g. robots). (Marciszewski, Stacewicz 2010).

5. NTMC models can provide theoretical basis for non-Turing models of mental activities – at least theoretical/formal ones.
Such models seem interesting because of: a) the possible **analogicity** of the brain (i.e. the biological basis of the mind), b) the **creativity** of the mind (which may require indeterminism

at the brain level), c) the ability of the mind to effectively use **infinite** objects (especially in mathematics).

References

Burgin, M., Dodig-Crnkovic, G. (2013). Typologies of Computation and Computational Models, arXiv:1312.2447 [cs].

Copeland, J. (2002). Hypercomputation. *Mind and Machines*, 12, 461–502.

Etesi, G., Nemeti, I. (2002). Non-Turing computations via Malament-Hogarth space-times. *International Journal of Theoretic Physics*, 41, 341–370.

Harel, D. (1987). *Algorithmics: The Spirit of Computing*, MA: Addison-Wesley.

Ord, T. (2006). The many forms of hypercomputation. *Applied Mathematics and Computation*, 178, 8–24.

Marciszewski W., Stacewicz, P. (2010). *Umysł – Komputer – Świat. O zagadce umysłu z informatycznego punktu widzenia*, Warszawa: Akademicka Oficyna Wydawnicza.

Shagrir, O. (2004). Super-tasks, Accelerating Turing Machines and Uncomputability. *Theoretical Computer Science*, 317, 105–114.

Shannon, C. (1941). Mathematical Theory of the Differential Analyzer. *J. Math. Phys. MIT*, 20, 337–354.

Stacewicz, P. (2018a). Czy informatykom musi wystarczyć nieskończoność potencjalna?. In R. Murawski, J. Woleński (Eds.), *Problemy filozofii matematyki i informatyki* (pp. 177-190). Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.

Stannett, M. (2003). Computation and Hypercomputation. *Minds and Machines*, 13, 115–153.

Turing, A. M. (1936). *On Computable Numbers, with an Application to the Entscheidungsproblem*. *Proc. Lond. Math. Soc*., 42, 230–265.

What differences can the body make? Remarks on morphological computation

Przemysław Nowakowski

According to Shapiro (2004), different body types determine different minds or different cognitive abilities. Wilson and Foglia (2011) expand this account and propose that bodies constrain, distribute, and regulate cognitive processing, therefore we can assume that different bodies should do this job in different ways. Here, I assume that if the computational approach to cognition is correct, and if the body really makes a difference, different bodies should determine differences in performing the computations underlying cognition. However, what differences can the body make?

Based on the contemporary literature (Caluwaerts et al. 2013; Ghazi-Zahedi et al. 2017; Miłkowski 2018; Müller, Hoffmann 2017; Nowakowski 2017; Seoane 2018), I will describe and assess the following options:

a. the body as *extending* resources for computation and storage;
b. the body as *increasing robustness* of performing the computations;
c. the body as *simplifying* the computations;
d. the body as *changing* the computations.

I will conclude my talk by objecting to approaches to morphological computation as a non-computation (Hewitson et al. 2018; Miłkowski 2018).

References:

Caluwaerts, K., D'Haene, M., Verstraeten, D., & Schrauwen, B. (2013). Locomotion without a brain: physical reservoir computing in tensegrity structures. Artificial life, 19(1), 35-66.
Ghazi-Zahedi, K., Langer, C., & Ay, N. (2017). Morphological computation: Synergy of body and brain. Entropy, 19(9), 456.
Hewitson, C. L., Kaplan, D. M., & Sutton, J. (2018). Yesterday the Earwig, Today Man, Tomorrow the Earwig?. *Comparative Cognition & Behavior Reviews*, *13*.
Miłkowski, M. (2018). Morphological Computation: Nothing but Physical Computation. *Entropy*, *20*(12), 942.
Müller, V. C., & Hoffmann, M. (2017). What is morphological computation? On how the body contributes to cognition and control. *Artificial life*, *23*(1), 1-24.
Nowakowski, P. R. (2017). Bodily processing: the role of morphological computation. Entropy, 19(7), 295.
Seoane, L. F. (2018). Evolutionary aspects of Reservoir Computing. *arXiv preprint arXiv:1810.04672*.
Shapiro, L. (2004) The Mind Incarnate; MIT Press: Cambridge, MA, USA.
Wilson, R.A.; Foglia, L. (2011/2016) Embodied Cognition. In The Stanford Encyclopedia of Philosophy; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA,

USA, 2016; Available online: https://plato.stanford.edu/archives/win2016/entries/embodied-cognition/ (accessed on 19 June 2017).

Conceptual structure of natural numbers
Modeling number cognition in conceptual spaces

Paula Quinon

People use natural numbers in different ways and for different purposes: for stating how many items there are in some collection or for rounding up a quantity, for enumerating elements from a collection, for computing (adding, multiplying, etc.), and for encoding (as an IT-routine or as a secret code). In most societies all these uses are handled by a single system of symbolic representations (*e.g.* the Arabic numerals). This multitude of meanings of the numerical expressions is responsible for the complexity of the process of acquiring the number concept: a child needs to acquire conceptual elements from many contexts and somehow put them all together in order to understand what numerical expressions mean and how one can use them. This is particularly difficult in our contemporary culture. As suggested by the archaeology of numerical systems (*e.g.* various notations in Maya's language) or testified by different numerical expression still used in certain spoken languages, the complexity of meaning encoded by a single digit or by a progression of numerals have increased.

We know today that this complexity of meaning has its roots in a multitude of cognitive factors amounting to the acquisition of number concept (Spelke 2000; Feigenson et al. 2004; Kinzler & Spelke 2007). For instance, humans, including infants, and also animals, have an ability to approximate quantities without counting (Dehaene 1997/2011). In addition to cognitive factors, symbolic representations of a given culture play a key role (Carey 2009). And the ability to recite number words without being able to count elements of a collection is different from the ability to choose instinctively a bag containing more candy. Number acquisition consists in correlating these different aspects into one unifying structure, the natural numbers.

The research of developmental psychologists shows that the conceptual content of numerical expressions increases gradually in children (Gelman & Galistel 1978; Sarnecka 2015). For instance, children learn the first names of numbers: "one", "two", "three" and "four" thanks to the ability to subitize, *i.e.,* visually determine the exact quantity of objects without counting – in order of magnitude and one-by-one. For example, a *two*-knower – usually a 2-year-old – knows what "one" and "two" mean, but associates random quantities to "three", "four", "five" etc.). It is also known that number concept development is not continuous. It is suggested that there is an "aha" moment when children (about 3.5yo) grasp a so-called "cardinality principle", which enables them to simultaneously understand meanings of bigger number names: "five", "six" etc. These – so called "*cardinality-principle*-knowers" – suddenly understand that the magnitude of a collection is expressed by the last name from the counting list they used to assess the cardinality of the collection. All these capacities – approximation, subitizing, counting list and cardinality principle - together amount to children's understanding of the number concept.

It is a theoretical challenge to understand how these different cognitive (approximation or subitizing) and cultural (counting list) factors combine and enable the number concept to be constructed in the mind of a child. In addition to its substantial scientific interest, a model of the process has important practical significance. An understanding of the process would make it possible to develop new pedagogical methods for helping children to understand the structure and use of natural numbers.

Most efforts to model the structure and development of number concepts have been based on computational models of the approximation process executed by Approximate Number System (ANS). These models aim at capturing the mathematical aspect of representations, how these representations differ depending on the magnitude of the input, depending on the type (progression or cardinality) of the input, modality of input, *etc.* They are oriented to disclose the algorithmic process in the background of human ability to approximate. For review see (Zorzi et al. 2005). In my talk, I will present an example of a connectionist model based on artificial neural networks formulated by Dehaene and Chagneux (1993).

I will then compare the computational models with models based on conceptual spaces. Conceptual spaces are a framework presented by Gärdenfors (2000, 2014). The main assumption of the theory of conceptual spaces states that meanings of words can be represented as a net of topological structures. At this point little is known about conceptual spaces for numbers, quantifiers and other concepts related to the different stages of number cognition. And hence, using conceptual spaces to model early number cognition, provides me with a particularly handy tool for making fine-grained conceptual distinctions. It is suited for modeling both symbolic and non-symbolic representations of knowledge and information.

Quinon and Gemel (2015; 2019) proposed a simplified conceptual space of pre-verbal ANS representations based on the idea that ANS-quantifiers are vague concepts. We used the conceptual spaces model of vagueness proposed by (Douven et al. 2013). The model we proposed ignores multiple general aspects of the behavior of ANS-related representations. We focused on representations that are activated in response to a constant visual input consisting of various sets composed of different quantities of blue dots distributed on a whiteboard. Formulating a model for a simple sensory input is a necessary step towards more comprehensive models accounting for other types of sensory inputs and also for symbolic input.

Moreover, conceptual spaces (Gärdenfors 2000; 2014) have been used to model the process of conceptual change in scientific theories (Gärdenfors & Zenker 2011; 2013; Zenker & Gärdenfors 2015). Following the analogy that children's thinking is similar to that of scientists (Gopnik 1996) and (Carey 1985), the next steps in the project will be using conceptual spaces for modeling the conceptual content that emerges at various stages of the number concept development (including subitizing and cardinality principle). We also want to use the models to show how these stages are related and how a child can possibly go from one stage to another (and what are the obstacles for such a transition).

Models of the conceptual structures of early number-theoretical (what are natural numbers?) and arithmetical knowledge (how to operate with them?) help us understanding how children construct numerical concepts and how it differs from the conceptual structure of the concept of natural numbers used by adults.

References

Carey, S.E. (1985). *Conceptual Change in Childhood.* MIT Press.

Dehaene, S. (1997/2011). *The Number Sense*. Oxford University Press.

Dehaene, S. Brannon, E. M., (2010) Space, time, and number: a Kantian research program, *Trends in Cognitive Science*, 2010 Dec; 14(12), 517-519.

Dehaene, S., & Changeux, J. P., (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience* 5, 390–407.

Douven, I., Decock, L., Dietz, R. & Égré, P. (2013). Vagueness: A Conceptual Spaces Approach, *Journal of Philosophical Logic* 42 (1): 137–160.

Feigenson, L., Dehaene, S., Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8(10), s. 307–314.

Gärdenfors, P. (2000). *Conceptual Spaces*. MIT Press.

Gärdenfors, P. (2014). *The Geometry of Meaning*, MIT Press.

Gärdenfors, P. & Zenker, F. (2013). Theory change as dimensional change: conceptual spaces applied to the dynamics of empirical theories. *Synthese* 190(6): 1039–1058.

Gelman, R. & Gallistel, C.R. (1978). *The child's understanding of number.* Harvard University Press.

Gemel, A., Quinon, P., (2015). The Approximate Numbers System and the treatment of vagueness in conceptual spaces, in: Gemel A., et. al. (eds.), Cognition, Meaning and Action, Jagiellonian-Lodz University Press, 87-108.

Gopnik, A. (1996). Scientist as a Child. *Phiosophy of Sciences* 63(4): 485-514.

Kinzler, K.D. & Spelke, E.S. (2007). Core systems in human cognition. In: von Hofsten & Rosander (eds.), *Progress in Brain Research* 164: 257-264.

Sarnecka, B.W. (2015) Learning to represent exact numbers In: Antonutti, M. & Quinon, P. (eds.), *Intensionality in Mathematics, Special Issue of Synthese*, Springer. Online First.

Spelke, E., (2000). Core Knowledge. *American Psychologist*. 55 (11): 1233–1243

Zenker, F. & Gärdenfors, P. (2015). Communication, Rationality, and Conceptual Changes in Scientific Theories. In: Zenker, F. & Gärdenfors, P. (eds.), *Applications of Conceptual Spaces The Case for Geometric Knowledge Representation*, Springer: 259–277.

Zorzi, M., Stoianov, I. & Umiltà, C. (2005). Computational Modeling of Numerical Cognition. In Campbell, J.I.D. (ed.) *Handbook of Mathematical Cognition*. Psychology Press: 67-84.

International Workshop on Computational Modelling
Warsaw University of Technology, May 14, 2019

Modeling ethical phronetic knowledge bases in A-Robots

Paweł Polak, Roman Krzanowski

The goal of social robotics design is to build robots that will interact with humans in a safe and ethical way. Such robots, or Artificial Moral Agents (AMA), would be expected to exhibit ethical capacities similar to that of humans. However, the conceptual and practical problems of building such agents have not yet been resolved. This paper proposes a new approach for developing ethical capacities in robots based on the concept of Aristotelian phronesis.

We outline one possible approach to the design of a phronetic robot. The proposed design uses a simplification of the original concept of phronesis (i.e., we implement a weak phronetic system), but it retains its characteristic features. In the proposed system, the ascent to the decision should be made through evaluating the existing use cases (UC) and comparing them to the specific situation, with the decision being made by selecting the most “ethically proximal” UC with the best outcome. The exact meaning of what the “ethically proximal case” is, needs a strict definition that can be operationalized. For instance, it is not clear What makes two or more UCs ethically similar or what makes a UC relevant to the situation.. The past UCs are not literal retentions of the past decisions, but rather abstracted situations. What is abstracted is not the material particularity of a situation but its ethical import. As vague as this statement is, it bars us from making decisions based on a simple “match” of the physical variables that describe the past cases.

We first need few definitions for the terms “agent”, “decision process”, “ethical fact”, and “ethical knowledge base”. An agent π obtains a set of observations O_{1i}, O_{2i}, …, O_{ni} about its environment (the state of the world). The index *i* denotes the specific set of observations (about the state-of-the-world); it may be the time index. An agent initiates the action a_i, resulting in an outcome q_i with a score s_i. We therefore have $p(O_{1i},O_{2i},\ldots O_{ni}\,)$Þ a_i and $a_i(q_i)$ Þs_i. F is a fact about a given situation (the state of the world in a given instance). This information is gathered by the agent from its environment. O_{1i} are specific instances of information. The collection of O_{1i} constitutes the fact (i.e., $F_i=(O_{1i},O_{2i},\ldots O_{ni})$). The use case (UC) comprises the state of the world, the action, the outcome, and the score, such that UC= $p(O_{1i},O_{2i},\ldots O_{ni}\,)$Þ a_i and $a_i(q_i)$ Þs_i. The Ethical Knowledge Base (EKB) is the set of UCs, such that EKB = [UC_1,…, UC_n].

The decision pathway for the autonomous moral agent (AMA) is depicted below.

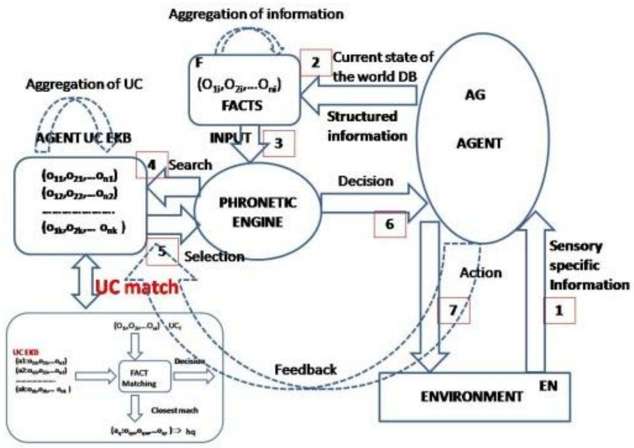

Figure 1 The decision pathway for an autonomous moral agent (AMA)

An Artificial Moral Agent (AMA) obtains information about the environment through its sensory systems (1), AMA processes this information into the proper format, and stores it in its database F of FACTS [$F_i=(O_{1i}, O_{2i}, \ldots, O_{ni})$**]** (2). The facts describe an agent's situation. When making a decision, the agent's inference engine (IE) takes facts that are relevant to the situation (3). The IE then searches the gent's EKB for UCs that are similar according to the specific proximity measure (4) and selects the decision that offers the most favorable outcome (5). The decision is passed back to the agent (6), which then acts in the environment (7).

The UC EKB contains the state of the world, actions, outcomes, and ethical scores. The results of actions are collected and feedback to the agent for integration in the UC EKB, as represented in Figure 1 by the dotted arrow. The presented schema of the functional blocks and decision pathways in the AMA with phronetic capacities is a high-level view that does not address the operational architecture of the system.

In the presentation we ask the following questions: what does constitute an ethical fact "FACTS [$F_i=(O_{1i}, O_{2i}, \ldots, O_{ni})$]" and how it can be formalized for the implementation in a digital system? How can the decision process be formalized "UC= $p(O_{1i},O_{2i},\ldots O_{ni}\ )$Þ a_i and $a_i(q_i)$ Þs_i." and what are its components and what is the format of EKB = [$UC_1, \ldots, UC_n$]? Several other issues with the implementations of AMA are mentioned, such as ethical proximity measures and ethical search strategies, aggregation of UCs in EKB.

Computational psychopathology. How computational modelling can contribute to the understanding of mental disorders?

Marcin Rządeczka

Undoubtedly, mental disorders can be counted amongst the most complex objects of scientific inquiry, which is not only a result of their multi-level casual hierarchy, mostly consisting of several to dozens of intrinsic and extrinsic factors, but also some very intricate criteria of diagnosis. It is nearly impossible to apply typological thinking to them, due to the fact that they form wide spectra of manifestations defeating any attempt to classify them as a discrete entity with a well-defined set of constituent symptoms. For the above-mentioned reasons, until recently, the science of psychopathology lacked any serious candidate for a unifying theoretical framework able to offer, at least, some draft explanation of the mental disorders' ultimate causes.

Indubitably, understanding mental disorders in terms of their proximate (i.e. mechanistic, neurobiological) causes is of great importance for both science and practice of medicine but offers no valuable answer to questions regarding the emergence of a certain disorder in a population. To answer such a question, one must delve deeply into the evolutionary biology and reframe the research perspective. Unsurprisingly, both populational and phylogenetical approaches shed new light on the origin of mental disorders, describing them as unavoidable failures of complex systems.

The fruitful union of computational biology and psychopathology aims to use computational modelling for the sake of creating the unified large-scale picture of mental disorders. According to so-called null-hypothesis, any sufficiently complex process influenced by, at least, several genetic and environmental factors, each of which in a highly variable manner, will be manifested as a broad spectrum of phenotypes, described roughly by the bell-shaped curve. In other words, this model predicts that abnormal cognition and behavior occurs in any population of organisms by default, and needs no special explanation at all. Low-frequency abnormalities (around 1%-3%) are the natural result of high variability of a functional phenotype. Bipolar disorder, schizophrenia and autism, all fall into this category.

Other interesting model includes viewing mental disorders as the by-product of the bodily behavioral defense system. Computational models of the human immune system provide some promising preliminary results about the possible costs and benefits of supplementing the innate and adaptive immunity with a set of non-specific avoidance mechanisms based on the smoke-detector principle. It can be applied to better understand specific phobias and OCD (Obsessive-Compulsive Disorder), to name just a few examples.

Some mental disorders can also be modelled as diseases of homeostasis and mismatch, resulting from the dysregulation of set points, which were primarily adjusted by natural selection to match the ancestral environment, where daily tribulations were radically different from the challenges of modern life. Interesting examples being depression and some eating disorders.

References

1) Giudice, M., (2018), Evolutionary Psychopathology. A Unified Approach.
2) Érdi, P., Bhattacharya, B., Cochran, A. (2017), Computational Neurology and Psychiatry.
3) Wallace, R. (2017), Computational Psychiatry. A Systems BiologyApproach to the Epigenetics of Mental Disorders.
4) Shackelford, T.K., Zeigler-Hill, V., (2017), The Evolution of Psychopathology.
5) Brune, M., (2016), Textbook of Evolutionary Psychiatry and Psychosomatic Medicine. The Origins of Psychopathology.