Beyond the Zermelo-Fraenkel Axiomatic System: BSDT Primary Language and its Perspective Applications

Petro Gopych
Universal Power Systems USA-Ukraine LLC
Kharkiv, Ukraine
pmgopych@gmail.com, pmg@kharkov.com

***Abstract*—A formalization of the recently proposed infinity hypothesis implying the common coevolution of the universe, life, mind, language, and society gives a possibility to introduce strict definitions of meaning and subjectivity, which spread beyond the traditional mathematics. This hypothesis leads to a semantic mathematics that is an implementation of the von Neumann's idea of a low-level "primary language" (PL). In this paper the formalization of this infinity hypothesis is further developed and some of its consequences are considered. In particular, a phenomenology formalization, definite and conditional meanings of the PL's words, their meaning complexity, categories and subcategories (hierarchies) of meaningful words, a way of the presentation of real numbers, the Cantor's continuum hypothesis, the PL's continuity-discreteness unity and uncertainty, non-Gödelian arithmetization by natural numbers and its relation to Chaitin's Omega-numbers, convention on truth, meaning ambiguity of words of different meaning complexity and its relation to Burali-Forti paradox are discussed. A validation of the PL is given. Some examples of meaningful computations using the technique of recently developed binary signal detection theory (BSDT) and the BSDT PL's perspective applications to solving the problems concerning the brain, mind and their faculties are briefly considered. It is emphasized super-Turing computations are typical for the BSDT PL as well as animal and human regular everyday meaningful communication.**

Keywords-infinity; meaning; subjectivity; phenomenology; context; categorization; attention; randomness; complexity; continuity-discreteness unity and uncertainty; arithmetization; continuum hypothesis; super-Turing computations.

I. INTRODUCTION

Recently proposed new infinity hypothesis [1] provides a possibility to strictly define such basic properties of mind as meaning and subjectivity. This hypothesis, contrary to the belief of some mathematicians [2], favors the view of mathematics as an invention of the mind. But mathematics is not only a product but also an instrument of the mind needed by humans to symbolically describe the world to better adapt to it. Mathematics may only be required and may only become possible in socially developed groups whose members are able to cooperate by means of a rather complex symbolic communication system or, in other words, by a language. In fact mathematics is an intrinsic part of the natural language (its fraction of maximal certainty) and can not be considered as something unrelated to it. Hence, mathematics as well as language is eventually the product of a particular human society and its culture, e.g., [3].

While humans are directing their efforts to mathematical problems which are in essence *external* with respect to their minds and faculties of minds (language, intuition, creativity, sociality, etc.), mathematics may be conceived, developed, and successfully applied in a completely *formal* way, ignoring the fact that it is inseparable from the mind/ meaning/subjectivity – it has been the course of the development of mathematics during thousands of years of its history. Rather recently this history was culminated in the design of formal axiomatic systems for mathematics as a whole – a finite number of most basic statements or *axioms* from which all mathematical theorems (correct assertions) can be derived in a finite number of logical inferences. This approach is known as a "finitist" one. The most famous and practically important of axiomatic systems is the Zermelo-Fraenkel (ZF) axiomatic system with the axiom of Choice (ZFC) [4]. The ZFC is widely recognized as a "standard" or "traditional" basis for all the contemporary mathematical formalism – a technique of writing out the axioms, theorems, and inference rules in a symbolic way. The very idea of the axiomatization and the goal of the famous David Hilbert's program [5] are to exclude from mathematics even the smallest traces of the mind/subjectivity, to reduce in this way mind-related ambiguities and to ensure as a result the highest possible (in the ideal case "absolute") logical rigor of it. But, as was demonstrated by Kurt Gödel [6], this enormous goal can never be achieved: in fact no finitist axiomatic system exists that leads to mathematical formalism that would simultaneously be consistent (all of its theorems do not contain logical contradictions) and complete (all of its theorems can be proved or, in other words, formally derived from the axioms). For the Hilbert's program and the whole axiomatic approach, this Gödel's incompleteness plays a destructive role. In spite of that, all theorems already proved and those that would only be proved in the future within the framework of the ZF/ZFC formalism remain valid while we are interested in problems that are not directed to our minds. An overwhelming majority of mathematical problems were till now exactly of this kind and, consequently, Gödel's incompleteness exerts no effect on them. Computational abilities of mathematical formalism that is based on ZFC-like axioms and ignores the mind were revealed and implemented by Alan Turing [7] as his famous abstract Turing machines.

The situation changes drastically as soon as humans start to address problems which are in essence *internal* with respect to their minds and faculties of minds. Now the minds have to symbolically describe themselves by means of methods created by them in a way that is comprehensible to other minds. Under such circumstances of self-reference, self-representation and self-awareness, standard (ignoring the mind) mathematics does not work and the impossibility of describing the mind by methods ignoring the mind manifests itself in some notorious paradoxes [4] and in Gödel's incompleteness [6] which becomes immediately practically significant. But even the most dogmatic formalists can not completely exclude the mind from their theories because the need always remains to explain (interpret) their special symbols in words of a natural language that names our elementary subjective experiences ("primitive intuitive knowledge"). As a result, mind-related meaning variability of natural languages penetrates into the mathematical formalism making it vague and insufficiently rigor. Henri Poincaré [8] formulated this problem as his "vicious circle argument" drawing the attention to the fact that at least some basic notions of known axiomatic systems are defined one through the other and, consequently, can not be treated as genuinely fundamental. Emphasizing the intuitionist aspects of knowledge, he also enunciated the intuitive origin of the fundamental in mathematics principle of induction. To solve these problems at least partially, formalists gather colloquial-word explanations of their symbols together, dub resulting collection the metamathematics, e.g., [9] and consider it as something different from the formalism itself.

There is also technically unpopular but methodologically important branch of mathematics known as L. E. J. Brouwer's intuitionism, e.g., [10]. It states that certain principal mathematical concepts (and, consequently, axioms as their symbolic representations) are immediately given to humans by their intuitions, though these concepts/axioms can never precisely be completed because over time they could be changed by further intuitions. Hence, the intuitionism and the formalism are axiomatic theories but with axioms introduced in different though similar ways (note, metamathematics, a part of the formalism, actually contains some elements of the intuitionism). It was John Lucas who first soundly stated [11] that the Gödel's incompleteness theorem [6] does not allow formal finitist explanation of the mind and presenting it as a Turing machine. Later these ideas were further developed by Roger Penrose [12], [13]. He, in order to explain mind/ consciousness, has argued the need of appealing to a so-far unknown hypothetical physics known as a "correct quantum gravity" that would be responsible for the emergence of our subjective experiences. In any case, Turing methods are insufficient to ensure mind/mind-related computations and to do this something "super-Turing" is required.

One can see, standard mathematics based on the ZFC or ZFC-like axiomatic approach becomes insufficient if we need to explain the mind and mind-related human/animal faculties. Numerous unsuccessful attempts to achieve an adequate description/understanding of such phenomena as mind/consciousness, e.g., [14], language, e.g., [15] or (mathematical) symbolism, e.g., [16] show these problems are tied in a Gordian knot and none of them, taken separately, could not fully be solved. It indicates we need a new mathematics spreading *beyond* the ZFC and equally successful in describing the phenomena that are *external and internal* with respect to the mind. Since standard mathematics successfully describes the mind-external world, it has to be a part of the required new mathematics. Hence, we need such a generalization of the ZFC that additionally takes into account a fundamental property of the world that is missed by the ZFC but crucially important for the emergence and maintenance of the human mind and mind-related faculties. We hypothesize [1], [17] this fundamental property is *the infinity of common "in the past" coevolution of the universe, life, mind, language, and society* (cf. Edward Wilson's idea of the "gene-culture coevolution" [18], Humberto Maturana and Francisco Varela's autopoiesis theory [19], Lynn Margulis's evolutionary symbiosis [20], and psychosomatic nets by Candace Pert [21]). In the present paper it is explained in which way this idea of infinity can be implemented by methods that are *beyond* the ZFC and regular Turing computations. The generality of our initial thesis entails the need to also address some other problems of great generality namely the context, meaning, attention, subjectivity, categorization, randomness, complexity, etc. In spite of that, this work is about science and not philosophy because it addresses a new mathematics and its practical computations.

For the sake of completeness, it is also needed to point to the opposite view usually accepted in the field of machine consciousness: the ZFC and Turing machines are sufficient for modeling the mind and serious obstacles that hinder achieving this goal are caused by severe but within the existing framework solvable technical problems, e.g., [22] - [24].

The rest of this paper is structured as follows. In Sections II to IV the hypothesis of concurrent infinity, based on it phenomenology formalization, and an implementation of the von Neumann's idea of a "primary language" (PL) are considered. It is also explained in which way the PL and recent binary signal detection theory (BSDT) [25] are closely related, why the BSDT PL spreads beyond the ZF/ZFC and why it may be treated as mathematics of meaningful computations. In Section V some details of the BSDT PL formalism are described, including the formalization of the notions of meaning, subjectivity and meaning complexity. Sections VI and VII describe a non-Gödelian arithmetization by natural numbers of all the BSDT PL expressions, their randomness and continuity-discreteness unity and uncertainty. The BSDT PL's convention on truth is considered in Section VIII. In Section IX the notion of conditional meaning is described and used to account for meaning ambiguity of BSDT PL words of different meaning complexities. A connection between the meaning ambiguity and Burali-Forti paradox is also demonstrated. Section X presents some numerical and empirical validations of the BSDT PL, including the existence in animals/humans of mirror neuron systems that implement typical super-Turing computations and BSDT PL super-Turing computers. In

Section XI examples of practically important given the context meaningful computations and some BSDT PL perspective applications are briefly discussed. Section XII gives conclusions.

II. BSDT Primary Language as von Neumann's Primary Language

It was John von Neumann who was perhaps the first mathematician claiming the need of another mathematics for brain computations. This article is an attempt of an implementation of his idea of a low-level "primary language *truly* used by the central nervous system," and structurally "essentially different of those languages to which our common experience refer" [26, p. 92]. It is this primary language that is this new "essentially different" (we suppose, spreading beyond the ZFC) mathematics for describing the mind and doing mind/brain computations.

Since the PL is a low-level language for a nervous system's internal computations, its symbolism should be relevant to the usual style of signaling in nerve tissues of animals/humans by means of short electrical impulses of given amplitude often called "action potentials" or "spikes". We assume informative messages of interest are conveyed and processed in the brain as *patterns* of such spikes. It is supposed, these patterns are represented in the PL as finite-dimensional binary spin-like (with components ±1) vectors distorted by a non-additive "replacing" binary noise [27]. The coding by binary noise and, as a result, using the "one-memory-trace-per-one-network" learning paradigm [28] are the main features of the BSDT [25] that gives *the best* coding/decoding rules for patterns of binary signals damaged by binary noise. Complete description of all non-discrete properties of neurons as, e.g., their electric-chemical interactions or refractory periods is included into the infinite context giving the meaning to a particular pattern of spikes or respective binary vector (Section V B). They also contribute to uncertainties discussed in Sections VI B and X E. BSDT +1/–1 code can not be replaced by the 1/0 code traditionally used in most computers. It is superior because binary +1/–1, ternary +1/0/–1, and quaternary "colored" +1/–1 codes naturally describe 1) neuron assemblies in different states of their synchrony and 2) different ways of reciprocal ("phase") transitions between them [27].

To formalize the well-known vision of the brain as a *selectional* device, e.g., [29] within the framework of the BSDT, by analogy with Turing machines, we have introduced the abstract selectional machines, BSDT ASMs [30]. ASMs ensure *the best* BSDT decoding and give a technical implementation of the idea that *meaning* of a finite symbolic message is mainly defined by its *infinite context*. The BSDT and its ASMs became also the ground for the BSDT neural network assembly memory model, NNAMM [31], and BSDT atom of consciousness model, AOCM [32]. They employ explicitly the idea of the equivalence between the meaning of a message and *subjective experience* or *primary thought* of an organism (perceiving agent) recognizing this message. Such an approach inevitably requires an extending of standard mathematics beyond the ZFC, to ensure the consideration of the phenomenon of meaning/subjectivity/privacy at mathematical level of logical rigor. The latter is the mandatory prerequisite for solving what is called the "hard" problem of consciousness [33].

At the same time the PL would remain an empty enterprise if there is no technique implementing it computationally. Fortunately, the BSDT gives such *the best* technique that is completely ready to be used. What is additionally needed to ensure its success is a methodology of its application to particular PL-specific problems, see Section XI. Hence, for the PL, the BSDT plays a two-fold role: on the one hand, it contributes to its substantiation; on the other hand, it gives its computational implementation. Consequently, it is natural to refer to the PL we propose as the BSDT PL. The PL in turn gives the BSDT the significance of the best technique of the PL's meaningful computations (e.g., Sections IV, V, X, and XI).

In the regular sense of this term, proofs are understood as finite sequences of formal symbolic logical transformations that draw the theorems from axioms. For BSDT PL statements, such formal proofs have strictly speaking no sense because, in this case, we are always interested in their meanings but standard mathematical formalism rejects meanings by definition. For the substantiation of BSDT PL statements, we will give neither theorems nor proofs. We will provide instead their unambiguous *constructions*, given our new premises (Section V) and known theorems of standard mathematics. Such a style of writing is "constructive" rather than formal.

In order to arrive at the BSDT PL we have mainly been motivated by biological and mathematical reasons. For this reason, along all this paper we will focus on those problems of life, mind, language and society that standard mathematics fails to resolve. The most acute of them and most amenable for the first BSDT PL application are perhaps the reliable language communication *without syntax* in humans and communication *without any language* at all in human infants and animals of the same or relative species. These problems are of great importance for linguistics, cognitive sciences, and artificial intelligence because their study informs us about the dynamics of language as a population phenomenon, bodily forms of signaling, and about a cognitive and bodily infrastructure for social interaction [34]. We also highlight the ranges of BSDT PL applications and show where and in which way it could be reduced to the reining standard mathematics.

The BSDT PL is of course not restricted to biology; it may also be useful, e.g., in physics but this direction of BSDT PL perspective applications remains out of the scope of this work.

III. Hypothesis of Concurrent Infinity, eBSDT and BSDT PL Phenomenology

The ZFC axiom of infinity postulates the infinity of the number of those elements/individuals that are used in ZFC theory of sets for the construction of these sets [4]. The meaning of the term "element/individual" is not specified in any way but it would be reasonable to believe (or at least one would prefer to believe) this infinity axiom reflects in a sense the tacitly assumed infinite richness of the world in which we live in, though what is "the world" is again explicitly not specified. In spite of obscure terminology used we have to agree that the ZFC seems to imply the infinite versatility of the world but certainly does not inherently imply the possibility of its evolution and development. The ZFC world is a stationary one. This theory allows the allocation of different aggregates of elements (the world's "currently visible" fragments) but does not allow any changes of neither the world as a whole nor its currently visible parts. The fragments of ZFC world (sets and subsets of elements) are "tautologically" [4] related to each other like ZFC theorems/tautologies that could be transformed one into the other with the help of simple or intricate but always *reversible* (as they produce the *tautologies* only) formal rules. The irreversibility of known irreversible functions originates from a randomness of processes, e.g., [27] they represent but are not from the ZFC. If one prefers to keep the elements as abstract and not related to the world entities then ZFC mathematics remains meaningless. In spite of that its computations may gain different meanings from different, e.g., physical problems they describe (see Section XI B).

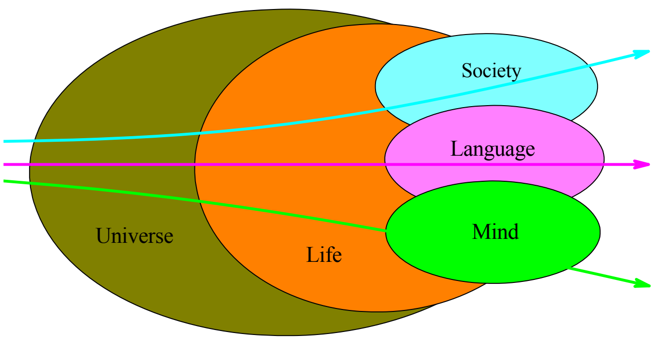

Figure 1. Hypothesis of concurrent infinity. Reciprocal relations between the universe, life, mind, language and society are shown as overlapping ovals of different colors. Arrows designate the course of their common infinite "in the past" (on the left) and open-ended "in the furure" (on the right) co-evolution.

The BSDT PL infinity hypothesis we are inaugurating is needed to introduce in mathematics the idea of permanent or "eternal" open-ended evolution and development of our (physical) world including animals/humans and their minds as a part of it. *In addition* to the ZFC-like infinite richness of the world, we postulate the infinity of common "in the past" and open-ended "in the future" co-evolution (Figure 1) of the universe, life, mind, language and society [1], [17], [32]. We also equate a *real-world physical device* devoted to the recognition of particular meaningful symbolic (binary for certainty) message originated from a thing of the world, complete binary *infinite on a semi-axis description* of the story of designing this device in the course of its infinitely long evolution from "the beginning of the world" until now, and *the meaning* of the message under consideration or *primary thought* it conveys. In addition *meanings* are interpreted as *subjective/first-person/private experiences* or respective *feelings* (*qualia*) and vice versa [1], [17], [32]. It is assumed, the BSDT PL world (it coincides with our physical world), is the total collection of what we call things, i.e., any inanimate objects, animate beings, and any relationships between/within them. All the world's things are permanently evolving, in the course of their common infinitely long *coevolution*, "in parallel" or *concurrently*, physically interacting with each other either directly or by their contributions to their common environment. To emphasize this issue we call our hypothesis *the hypothesis of concurrent infinity*. The BSDT extended by this hypothesis is referred to as *the eBSDT*; it is the basis for the BSDT PL we describe here as well as BSDT AOCM [32].

Our infinity hypothesis defines simultaneously a phenomenology (Figure 2), that is, explicit relationships between human subjective experiences and real world things humans perceive. The phenomenology is the branch of philosophy and science that emphasizes the role of/ concerns mainly with perceptual/subjective aspects of our knowledge and our minds. In its present form, it was established and strongly advocated by Edmund Husserl, e.g., [35]. For the recognition of the meaningful message of interest (a binary vector x_j^i given its infinite binary context c_{xi}, Section V), our phenomenology postulates the use of a real-world implementation of the BSDT ASM devoted to process the x_j^i ([30] and box 1 in Figure 2) and, consequently, it is *the BSDT PL phenomenology*. It not only connects our feelings to things we perceive (boxes 1 to 3, some details see in [32]) but ensures also their *formalization* giving a possibility (box 4) of establishing a set of formal rules defining in which way to deal with meanings and feelings in terms of standard mathematics (Section V). From a common-sense point of view, taken together, explicit phenomenological traits of our infinity hypothesis do give the theories based on them (BSDT PL and BSDT AOCM [32]) a little taste of "strangeness".

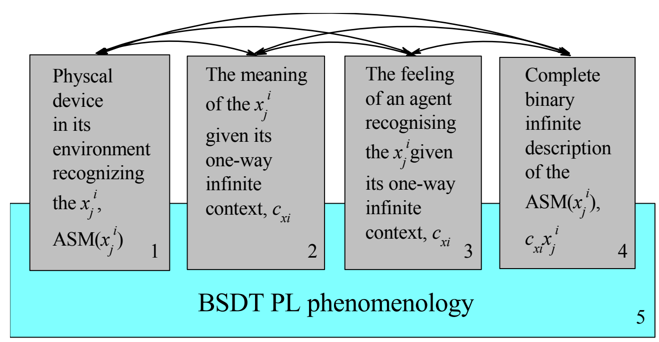

Figure 2. The BSDT PL phenomenology. Real-world physical devices recognizing the strings/vectors x_j^i and their equivalents in meaning, feeling, and symbolism domains (boxes 1 to 4) give rise together to BSDT PL phenomenology (box 5). Bidirectional arrows are used to designate the signs of equivalence and "paradigm shifts" or transitions "between incommensurables" [36, p. 150].

The distinctive feature of the hypothesis of concurrent infinity, which is rather difficult to comprehend, is that it *literary equates* usually *incommensurable* entities, related to quite different domains – real-world physical devices (box 1 in Figure 2), meanings of names of real-world things (box 2), subjective feelings (box 3), and infinite strings of symbols (box 4). In particular, our suggestion (boxes 1 and

2) that the meaning of a thing's name is identical to the physical device recognizing this name but not to the thing itself seems at the first glance counterintuitive. At least Ludwig Wittgenstein stated the opposite: "A name means an object. The object is its meaning." [37, 3.203]. I.e., he equated meanings of words and *the things* to be named while we equate meanings of words and *an agent's devices* recognizing these words. In other words, it is assumed the meaning of a thing's name is a specific internal activity or *psychological state* or *primary thought* of an organism (sensory agent) perceiving this thing. Name meaning is the property (momentary internal state) of a perceiving agent and not the property (feature, trait, state) of the thing to be named. It is easier to intuitively acquire this statement, if to remember that the things of the world are always given to us *indirectly*, through our sensory organs and respective patterns of sensory signals in our nerve tissues.

IV. The Idea of Phenpmenology Formalization and Semantic Mathematics

To do the formalization of BSDT PL phenomenology, we invoke our suggestion that an agent's psychological states, name meanings, primary thoughts and physical devices devoted to recognize the names are equivalent to infinite on a semi-axis binary strings that share their infinite on a semi-axis initial parts/beginnings or, in other words, that have common *prehistory*. It is a *pre*history and not the history because it describes the beginning of everything in the world and remains always essentially *unspecified*. We know about the prehistory that its existence is one-way infinite and common for all the things of the world but *nothing* more. Of this follows that a particular *infinite on a semi-axis* binary string describing the meaning of a thing's name (the physical device that recognizes this name and feeling the thing causes in a perceiving agent) has *infinite on a semi-axis* beginning that coincides bit-by-bit with *infinite on a semi-axis* beginnings of other *infinite on a semi-axis* binary strings describing the meanings of names of other things of the world. Consequently, to formalize the operations with name meanings and respective feelings (Section V), it suffices to fix such an arrangement of their one-way infinite meaning descriptions when their infinite on a semi-axis beginnings (initial parts or prehistory) coincide completely and, after excluding these common beginnings from the consideration, to deal with their *finite-in-length string remnants* only by methods of standard mathematics. Such mathematics of computations with *meaningful* one-way infinite and specially arranged strings we call *meaningful* or *semantic mathematics*.

The fact that meaningful/semantic computations with finite binary strings are defined given their common infinitely long on a semi-axis prehistory or under condition that their infinite on a semi-axis beginnings coincide bit-by-bit completely transforms them into a kind of *conditional computations*. Hence, BSDT PL computations obeying the demands of the hypothesis of concurrent infinity and respective BSDT PL phenomenology are ultimately the computations performed by methods of standard mathematics *given additional boundary conditions specified by a binary text that occupies completely an infinite semi-axis*. The BSDT PL as a semantic mathematics is *a generalization of standard mathematics* for the case of operations with one-way infinite binary strings having common one-way infinite beginnings and, simultaneously, a kind of *standard mathematics conditioned by* infinitely large amounts of additional assumptions written as an infinite on a semi-axis binary string. Once these additional boundary conditions (assumptions) are discarded, the mathematics of meaningful computations disappears and becomes the standard ZFC mathematics that by definition ignores meanings of its theorems/computations.

Since meaningful (semantic) computations are dealing with infinitely long strings/messages (i.e., taken as a whole genuine real numbers), they cannot be performed by regular Turing machines. To cope with semantic computations, a *super-Turing* computational technique and its implementation in the form of a physically constructible super-Turing computer [38] are obviously required (see Section X B and D). Like Turing machines implement the computations in standard ZFC or ZFC-like mathematics, super-Turing machines should implement the computations in semantic mathematics

V. Elements of the BSDT PL Formalism

The acceptance of the hypothesis of concurrent infinity transforms ZFC mathematics into the BSDT PL whose formalism differs to an extent from what we customary use.

A. Alphabet of Meaningful Words

As the BSDT PL is a kind of standard mathematics, the alphabet of the latter may be accepted as the alphabet of the former, with reservations concerning the specificity of semantic mathematics. The most important of them is that basic BSDT PL objects are *infinite* on a semi-axis symbolic (binary for certainty) meaningful strings that have common coinciding bit-by-bit *infinite* on a semi-axis meaningful beginnings. The other is that all these strings are written in the BSDT format as spin-like ±1-sequences and their finite-in-length end-fractions are processed by respective BSDT ASMs. In other words, it is assumed, all BSDT PL finite binary strings are coded using replacing binary noise [27] and decoded by BSDT ASMs [30] that give in this case the best decoding rules. The type of coding of infinite-in-length but explicitly unspecified common beginnings of meaningful strings does not matter. Meaningfulness of BSDT PL one-way infinite strings (and their equivalence to real-brain physical devices) is actually *postulated* by the hypothesis of concurrent infinity (Sections I and III, Figure 1) and respective BSDT PL phenomenology (Sections III and IV, Figure 2).

B. Vocabulary of Meaningful Words

The distinctive feature of the BSDT PL is that its basic elements (words) are meaningful. Their meanings are introduced as follows.

1) Meaningful simple words: Let us *arbitrary* choose one of BSDT PL one-way infinite binary strings, e.g., the c_{x0}

as a "master string". The length of it in bits, $l(c_{x0})$, equals by definition $\aleph_0$: $l(c_{x0}) = \aleph_0$ where $\aleph_0$ is the Georg Cantor's aleph-nought. If to divide the c_{x0} in the ith *randomly* chosen place into two parts then its finite and infinite fractions could respectively be thought of as an i-bits-in-length *simple meaningful word* x_j^i naming the ijth thing of the world and *the context* c_{xi} in which this word appears. Taken separately, x_j^i is meaningless. Resulting string $c_{x0} = c_{xi}x_j^i$ may be treated as the jth value of the string function/form $C(x^i) = c_{xi}x^i$ where string variable x^i is a string template of i empty cells needed to produce the strings x_j^i by filling this cell template in different i-length arrangements of +1s and −1s ($c_{xi}x^i$ is a concatenation of infinite binary string c_{xi} and cell template x^i). At a given value of i, with the help of the $C(x^i)$, 2^i of different strings $c_{xi}x_j^i$ can be generated with the same context c_{xi} and different affixes x_j^i. An affix x_j^i may simultaneously be treated as the ijth i-bits-in-length binary string, message, computer code/algorithm, vector or point in the space S_{xi} ($x_j^i \in S_{xi}$; j = 1, 2, … 2^i), element of the set S_{xi} of the cardinality $|S_{xi}| = 2^i$, BSDT PL word or name (indices i and j point to a particular thing of the world, see Section V B3). Depending on the current context, these terms will further be used interchangeably.

By changing the values of i from zero to infinity, the function $C(x^i)$ allows to generate (construct) any finite binary string x_j^i of any length $i = l(x_j^i)$ given its infinitely long context c_{xi}. All resulting strings $c_{xi}x_j^i$ constitute together an *ultimate or proper class* S_{cx0} – the set that is not a member of any other set [39], $c_{xi}x_j^i \in S_{cx0}$. The term "proper class" may intuitively be interpreted "as an accumulation of objects which must always remain in a state of development" [40, p. 325]. The items of the S_{cx0}, $c_{xi}x_j^i$, are uniquely specified by their i-bits-in-length affixes (right-most end-fractions) x_j^i. The c_{xi} is common infinite context for all the x_j^i of the length i that have different arrangements of their ±1 components; i = 0, 1, 2, … and j = 1, 2, …, 2^i. The principal property of elements of an S_{cx0} is that, in the sense of Cantor, they are all of the same infinite length $\aleph_0$ (are countable) but, in spite of that, they and their infinite fractions are explicitly *comparable* and may be a number of bits longer or shorter with respect to each other. Of the vantage of standard mathematics, the latter conclusion is fundamentally impossible though it is the norm for the BSDT PL due to the common infinite beginning of all its one-way infinite strings. For example, if meaningful simple words x_j^i and x_j^k obtained with the help of forms $C(x^i)$ and $C(x^k)$ given the same master string $c_{x0} \in S_{cx0}$ are of different lengths (e.g., $k > i$) then $l(c_{x0}) = l(c_{xi}x_j^i) = l(c_{xk}x_j^k) = l(c_{xi}) = l(c_{xk}) = \aleph_0$ but $l(c_{xi}x_j^i) - l(c_{xi}) = i$, $l(c_{xk}x_j^k) - l(c_{xk}) = k$, $l(c_{xi}x_j^i) - l(c_{xk}x_j^k) = 0$, and $l(c_{xi}) - l(c_{xk}) = k - i > 0$ (for $c_{xi}x_j^i$ and $c_{xk}x_j^k$, their the largest common infinite beginning is c_{xk}; see Figure 3).

Note, infinite words [41] of automata theory have no common infinite beginnings and remain within the framework of traditional mathematics

2) Meaningful composite words/sentences and their focal and fringe constituents: If string variable x^i consists of variables u^p and v^q then $x^i = u^pv^q$ with $i = p + q$; u^pv^q is a concatenation of cell templates u^p and v^q. The values of variables x^i, u^p, and v^q are respectively the strings x_j^i, u_r^p, and v_s^q that are the members of sets S_{xi}, S_{up}, and S_{vq} whose cardinalities are respectively $|S_{xi}| = 2^i$, $|S_{up}| = 2^p$, and $|S_{vq}| = 2^q$; $S_{up} \subseteq S_{xi}$ and $S_{vq} \subseteq S_{xi}$. The values of composite variable $x^i = u^pv^q$ are the composite words $x_j^i = u_r^pv_s^q$, the order of the composite word's constituents is essential for them (see an example in Figure 4). Composite variables consisting of any number of their constituents may similar be constructed. Composite space S_{xi} may also be interpreted as either the S_{up} whose vectors are colored in 2^q colors or the S_{vq} whose vectors are colored in 2^p colors. If so, then p and q are the measures of discrete "colored" non-localities of vectors in spaces S_{vq} and S_{up}, respectively [17], [27]. Similar colored (blue-and-red) binary spaces (three-dimensional "colored Boolean cubes") have earlier been used for representing the Boolean functions of one-dimensional cell automata, e.g., [42, ch. 6]. The rainbow of colors in finite-dimensional binary spaces discussed here is a direct generalization of two-color spaces of any dimensionality we previously introduced [27] to describe the coding of signals in nerve tissues of animals/humans.

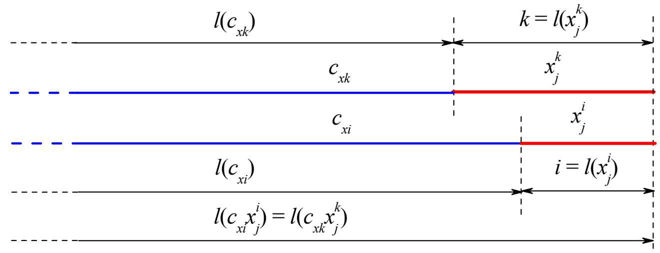

Figure 3. Meaningful simple words x_j^i and x_j^k (red line segments), their infinite on a semi-axis contexts c_{xi} and c_{xk} (blue line segments), and their meanings $c_{xi}x_j^i$ and $c_{xk}x_j^k$ (red and blue line segments taken together). The lengths of red line segment in bits, $i = l(x_j^i)$ and $k = l(x_j^k)$, are ensemble complexities of x_j^i and x_j^k ($k > i$), the lengths of blue line segments, $\aleph_0 = l(c_{xi}) = l(c_{xk})$, are their context complexities ($l(c_{xi}) - l(c_{xk}) = k - i > 0$); the lengths of red and blue line segments taken together, $\aleph_0 = l(c_{xi}x_j^i) = l(c_{xk}x_j^k)$, are their meaning complexities ($l(c_{xi}x_j^i) - l(c_{xk}x_j^k) = 0$, Section V C). Colored line segments denote the strings themselves, arrows designate their lengths. Dashed ending of lines on the left designate their infinity "in the past".

Isolated composite words are meaningless. Like simple words, they take their meanings from their infinite contexts and from themselves. For this reason, definite meanings have either whole composite words or their right-most fractions only. The right-most fraction of a composite meaningful word occupies an animal's dynamically created "focus of attention" and is called the "*focal*" word. The composite word's non-focal component is the focal word's "*fringe*" (by analogy with fringes of memory and consciousness [31], [32]) or the focal word's short-range or immediate or local context. If internal structure of a composite word is ignored then it is interpreted as a focal word that has zero-length fringe.

Meaningful composite words $x_j^i = u_r^pv_s^q$ are thought of as BSDT PL meaningful *sentences*. The focal word v_s^q corresponds to a sentence's feature/attribute that is currently in the focus of an animal's attention; its fringe u_r^p is the fringe of the animal's memory or consciousness. A composite word's "holophrasical" presentation (without noticing its internal structure), e.g., x_j^i corresponds to the perception/understanding of a sentence as a whole whereas its "analytical" presentation as, e.g., a set of possible focal words v_s^q with $1 \le q \le i$ gives the sentence's meaning as a series of meanings of its simple focal words. Composite word's holophrasical presentation describes the perception of a thing as a whole (*diffuse* focus of attention) whereas its analytical presentation describes its perception as a series of its attributes (*acute* focus of attention). Any paraphrase of BSDT PL sentences (any other choice of their constituents) cannot change their whole meanings and in that sense the BSDT PL lacks "compositional semantics" [43]. Owing to our infinity hypothesis (Figure 1) and its phenomenology (Figure 2), meaningful BSDT PL sentences (meaningful composite words) are simultaneously real-brain devices processing these sentences. For this reason, the number of a composite word's constituents may be treated as an animal's *logical or reasoning deepness*. Since logical or reasoning deepness measured in humans is 3 to 5 [44], the biologically most plausible number of components constituting BSDT PL composite words is expected to be of the same value, 3 to 5.

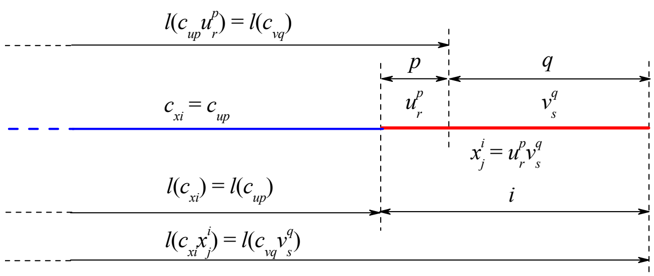

Figure 4. Meaningful composite word/sentence $x_j^i = u_r^p v_s^q$; v_s^q is a focal word of the sentence x_j^i, u_r^p is the focal word's fringe. Designations as in Figure 3.

As composite words are treated as BSDT PL sentences, the set of rules for the construction of meaningful composite words from a set of its possible constituents produces the BSDT PL *syntax*. Since any operations on simple and composite meaningful words should always be performed given their meanings (taking into account their common infinite beginnings), BSDT PL semantics (interpretations of words) is primary with respect to its syntax (rules for the construction of words) though they are of course closely related. If internal structure of meaningful composite words/sentences is ignored and they are perceived as a whole then communication with their help does not appeal to BSDT PL syntax and, consequently, it is performed *without syntax*, which is typical for animals and human infants [34].

3) Naming the things by meaningful words: All the meaningful words x_j^i constitute the BSDT PL *vocabulary* – a set of words that name, given their infinite context, all the things of the world, known as well as unknown but only conceivable. The number of things of the world is supposed to be infinite but countable, like the number of different meaningful strings $c_{xi}x_j^i$ related to a given proper class (Section VI A). BSDT PL vocabulary is always limited though, by request, may arbitrary be enlarged to the extent constrained mainly by particular animal's morphology only (new meaningful names may always be *constructed* and added to the vocabulary). Meanings of BSDT PL words are the ones that animals keep *actually* in their minds because, for an animal's survival, it is needed, its nervous system does not lie to itself. That is the reason why the BSDT PL should be successful as a truly primary language.

BSDT PL word x_j^i (i-bits-in-length sequence of +1s and –1s) is the jth pattern of spikes in the ith brain area equipped by the ijth BSDT ASM devoted to recognize the x_j^i, the name of the ijth thing of the world. This area may contain up to 2^i of such ASMs (cf. Figure 8). The reservation concerning brain areas (perceptual submodalities) is needed to connect the x_j^i to its context c_{xi} that specifies together with the x_j^i itself particular real-brain physical device recognizing the x_j^i and giving a meaning to it. Hence, BSDT PL meaningful words (patterns of +1s and –1s) are simultaneously the patterns of nerve impulses (spikes) in specific brain areas but certainly not the words of any of natural languages. With respect to our primary language natural languages are the *secondary* ones [26]. Invoking the notions of neuroscience (e.g., spikes or brain areas) for underpinning the theory's formal mathematical issues seems rather strange but does not reduce the theory's rigor. A reference to neuroscience is inevitably needed to give an explicit specification of one of the theory's principal paradigm shifts [36] shown by arrows in Figure 2, namely the shift from the domain of mathematical symbolism (box 4) to the domain of physically constructible brain devices (box 1) devoted to the recognition and processing of symbolically presented messages originated from things of the world (cf. Section X B and D).

C. *Meaning Complexity and Levels of Meaning Uncertainty of Meaningful Words*

The first thing that is needed to operate with meaningful words is a way of comparing them. Available methods do not hold in the case of concurrently infinite words.

1) The quest for a new infinity measure: All BSDT PL meaningful strings are one-way infinite, have an infinite length $\aleph_0$, and share their infinite beginning the length of which is again $\aleph_0$. Since their beginnings (distributed but precisely arranged and fixed "points of origin") are always the same (completely coincide), their end-points may take different locations and respective one-way infinite strings may in general be a number of bits longer or shorter with respect to each other. Figure 3 shows meaningful words whose complete string representations have the same end-points but whose string contexts have different end-points. Figure 5 illustrates another case: meaningful words whose complete string representations have different end-points but whose string contexts have the same end-points. We see the strings of the same infinite length in the sense of Cantor (that are countable) may be of different infinite length in the sense of the BSDT PL and, consequently, it is needed to introduce a measure of lengths of such infinite strings (i.e., a measure of infinity) that should quantify their total lengths and the distinctions in positions of their end-points.

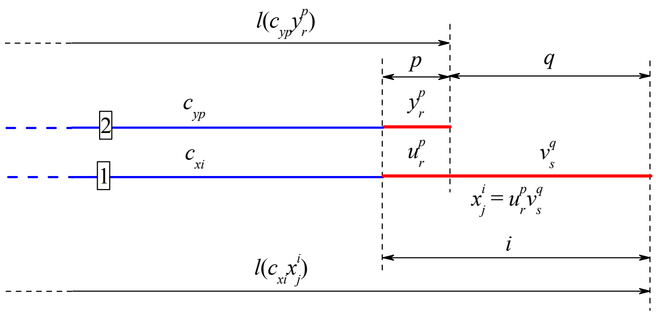

Figure 5. Meaningful words y_r^p and x_j^i of different in q bits meaning complexities, $l(c_{xi}x_j^i) - l(c_{up}y_r^p) = q$. Taken separately, strings 1 and 2 are of the same infinite length, $l(c_{xi}x_j^i) = l(c_{yp}y_r^p) = \aleph_0$. x_j^i and y_r^p are focal fractions of strings 1 and 2 and, consequently, are meaningful. If x_j^i is a composite word, $x_j^i = u_r^p v_s^q$, then u_r^p is a fringe of meaningful focal name v_s^q (in line 1 and 2, y_r^p and u_r^p may bit-by-bit coincide). Designations as in Figure 3.

2) Meaning complexity: We refer to the lengths $l(c_{xi}x_j^i)$ of infinite-on-a-semi-axis binary strings $c_{xi}x_j^i$ with their common infinite beginning (the context, c_{xi}) and their different explicitly specified affixes/meaningful *simple* words x_j^i as *meaning complexities* of these words (see Figures 3 to 5). If such strings have the same end-points then their affixes are understood as meaningful simple words of the same meaning complexity (Figure 3). If the end-points of such strings do not coincide then meaning complexities of respective meaningful words do not coincide too and the largest meaning complexity has the word whose meaningful string description has the end-point that is located to the right of end-points of other meaningful strings (in Figure 5, $c_{xi}x_j^i$ has larger meaning complexity than the $c_{yp}y_r^p$ because the difference $l(c_{xi}x_j^i) - l(c_{yp}y_r^p)$ equals $q > 0$). The length $l(c_{xi})$ of one-way infinite context string c_{xi} we call the *context complexity* of meaningful word x_j^i. For different meaningful simple words, their context complexities may be different (Figure 3) as well as the same (Figure 5). The length of the word x_j^i (it equals i bits) we call *ensemble* or *statistical* or *Boltzmann* or *Shannon complexity* of this word. To explain the latter, let us note that according to Claude Shannon [45], information or entropy of a set of 2^i of statistically independent binary messages x_j^i (the values of string variable x^i) is defined as $H(x^i) = -\sum_j \log_2(P(x_j^i))/|S_{xi}| = -2^i \log_2(1/2^i)/2^i = i$ where $P(x_j^i) = 1/|S_{xi}|$ and $|S_{xi}| = 2^i$ are respectively the probability of occurring of any of the x_j^i (they are here meaningless) and the total amount of different x_j^i, $j = 1, 2, \ldots, |S_{xi}|$.

Hence, meaning complexity of a meaningful simple word x_j^i equals the sum of its context complexity, $l(c_{xi})$, and its ensemble complexity, $l(x_j^i) = i$: $l(c_{xi}x_j^i) = l(c_{xi}) + i$ where the first item gives the length of complete description of the common part of the story of designing the devices devoted to the recognition of different meaningful words x_j^i and the second item gives the properties of any of the x_j^i averaged over the set of them, S_{xi}. String description $c_{xi}x_j^i$ of the story of designing the device devoted to recognize the x_j^i is actually *the shortest* evolutionary algorithm/instruction for such design and, consequently, the length of this algorithm/story is its *Kolmogorov* or *algorithmic complexity*, e.g., [46]. Of this follows, the notion of meaning complexity embraces the notions of Kolmogorov complexity/information and Shannon complexity/ information/entropy. Meaning complexity specifies the algorithm of designing a recognition device and reflects the complexity of this device dedicated to processing a particular meaningful word (an animal's respective internal/psychological state) and not the complexity of the thing named by this word.

3) Levels of meaning uncertainty: Let us now consider the case of meaningful *composite* words, e.g., $c_{xi}x_j^i$ with $x_j^i = u_r^p v_s^q$. If to *dynamically* fix p left-most components of an x_j^i as a particular u_r^p then $c_{xi}x_j^i = c_{xi}(u_r^p v_s^q) = (c_{up}u_r^p)v_s^q = c_{vq}v_s^q$ where $c_{xi} = c_{up}$, $c_{vq} = c_{up}u_r^p$, $x_j^i \in S_{xi}$, $u_r^p \in S_{up}$, and $v_s^q \in S_{vq}$ (see line 1 in Figure 5). Infinite strings $c_{up}u_r^p \in S_{cu0}$ and $c_{vq}v_s^q \in S_{cv0}$ are the members of ultimate classes S_{cu0} and S_{cv0} that are different because they are generated by master strings c_{u0} and $c_{v0} = c_{x0}$ that share their beginning but differ in length in q bits. Owing to our infinity hypothesis the lengths of strings $c_{up}u_r^p$ and $c_{xi}x_j^i = c_{vq}v_s^q$ are comparable and the former is $l(c_{xi}x_j^i) - l(c_{up}u_r^p) = i - p = q > 0$ bits shorter (has smaller meaning complexity) than the latter (note, proper classes S_{cv0} and S_{cx0} coincide and, consequently, $l(c_{xi}x_j^i) = l(c_{vq}v_s^q)$). Since they both have infinite contexts, infinite strings $c_{xi}x_j^i$ and $c_{up}u_r^p$ are meaningful (in Figure 5, $c_{xi} = c_{up}$). But x_j^i (a simple focal word) and v_s^q (a focal fraction of $x_j^i = u_r^p v_s^q$) have the *definite* meanings while u_r^p (a fringe of the v_s^q in $x_j^i = u_r^p v_s^q$) a *conditional* meaning (see also Section IX). The latter may be compared with definite meanings of meaningful focal words with 2^q-state uncertainty defined by colored 2^q-non-locality of fringe words u_r^p. We refer to words whose meanings may only conditionally be defined as words of certain *levels of meaning uncertainty*. For this reason, the level of words of definite meanings is postulated to be zero (e.g., x_j^i, y_r^p, and v_s^q in Figure 5) while the level of uncertainty of words having conditional meanings (e.g., u_r^p in Figure 5) is a positive integer $q = i - p > 0$. The level of meaning uncertainty specifies the fringe's position in the body of its composite word (q bits to the left of the end-point of the whole meaningful string) and simultaneously, the degree, 2^q, of its colored non-locality. If there are words of definite meanings of different meaning complexity, e.g., x_j^i and y_r^p in Figure 5 then the one that has larger meaning complexity, x_j^i, may have a fringe, u_r^p, that coincides bit-by-bit with the word of definite meaning of smaller meaning complexity, y_r^p. In spite of that the meanings of y_r^p and u_r^p are essentially different. Attempts of comparing definite meanings of words of different meaning complexities (e.g., y_r^p and v_s^q or x_j^i) also lead to meaning uncertainties we quantify by the level of uncertainty of the u_r^p coinciding bit-by-bit with the y_r^p, i.e., the q for the example in Figure 5 (see Section IX).

4) Relative measurements of infinity using meaning complexity and levels of meaning uncertainty: Meaning complexity and levels of meaning uncertainty of BSDT PL meaningful words are the parameters needed to ensure a *relative* comparison of lengths of one-way infinite strings sharing their infinite beginning and, as a result, the comparison of word meanings, definite as well as conditional. Traditional (in the sense of Cantor) comparison of lengths of such strings by counting the total amount of their bits has no sense here because resulting lengths are always the same and equal to $\aleph_0$ bits. Consequently, meaning complexity and levels of meaning uncertainty (as parameters specifying the infinity) exist in the framework of the BSDT PL only and have their roots in the hypothesis of concurrent infinity and the technique of proper classes. Meaning complexity embraces, given the context c_{xi}, Shannon-type ensemble complexity (the length of a word x_j^i in bits) specifying the word's ensemble properties (*averaged* over the ensemble of 2^i of x_j^i) and Kolmogorov-type algorithmic complexity (the length in bits of complete irreducible infinite evolutionary algorithm/instruction $c_{xi}x_j^i$ for designing the ASM that selects the meaningful x_j^i) specifying *the sameness* or *individual properties* of the device selecting the x_j^i and, through it only, *the sameness* or *individual properties* of the thing named by the x_j^i.

In this article, our meaning complexity is not compared with numerous other complexity definitions (the notion of the level of meaning uncertainty is new at all). We note only that most of them, to take into account the current actual context, attempt to estimate it, in one or another way, in a *finite* manner. For example, using a finite estimation of what is called an "effective complexity" (that is a loose counterpart to or a finite estimation of our context complexity), Murray Gell-Mann and Seth Lloyd [47] combine Kolmogorov complexity/information and Shannon complexity/information into a finite "total information." Hence, the meaning complexity's crucial distinction is *the genuine explicit infinity* of its descriptions of meaningful words/sentences – the faculty that is fundamentally impossible within the framework of standard ZFC or ZFC-like mathematics.

D. Categories and Subcategories (Ontologies, Hierarchies) of Meaningful Words, Semantic Rule of Identity, Randomness and Irreducibility of Synonyms

Meaningful words could be organized in structures that are themselves meaningful and have rich properties.

1) Categories and subcategories: The values $c_{xi}x_j^i$ of the form $C(x^i) = c_{xi}x^i$ define *a category* (notion, concept) of 2^i of meaningful words x_j^i that are here called *synonyms*. Meanings of these synonyms are given by the strings $c_{xi}x_j^i$, $x_j^i \in S_{xi}$ and $|S_{xi}| = 2^i$. Considering the x^i as a composite variable, $x^i = u^p v^q$, allows (given the context c_{xi}) a sub-categorization of items of the category $C(x^i) = C(u^p v^q)$. If to fix the context $c_{xi} = c_{up}$ and a value of u^p, e.g., u_r^p ($u_r^p \in S_{up}$ and $|S_{up}| = 2^p$) then we obtain the category's *the* pr*th subcategory* $C_{pr}(v^q) = (c_{up}u_r^p)v^q = c_{vq}v^q$ of synonyms v_s^q of definite meanings $(c_{up}u_r^p)v_s^q = c_{vq}v_s^q$ where $v_s^q \in S_{vq}$ and $|S_{vq}| = 2^q$. If to fix the context $c_{xi} = c_{up}$ only then we obtain the category's the qth subcategory $C_q(u^p)$ of 2^p of synonyms u_r^p of conditional meanings of q-level meaning uncertainty (strings u_r^p occupy fringe positions in composite words $x_j^i = u_r^p v_s^q$, $q = i - p$). The number of subcategories $C_{pr}(v^q)$ whose synonyms have definite meanings equals the number of synonyms u_r^p of the subcategory $C_q(u^p)$ whose items have conditional meanings, i.e., 2^p; for the considered case of two-component composite words, the number of sub-categories whose synonyms have conditional meanings equals 1. If $p = 0$, the category $C(x^i) = C(u^p v^q)$ may be thought of as its own subcategory whose focal words $v_s^q = x_j^i$ ($q = i$) have common zero-length fringe u_0^0, $C(x^i) = C(u^p v^q) = C_{00}(v^q)$; meanings of synonyms of the $C_{00}(v^q)$ are given by the strings $c_{xi}x_j^i = c_{u0}u_0^0 v_s^q = c_{vq}v_s^q$ where $c_{xi} = c_{vq} = c_{u0}u_0^0$. For the case $i = 3$, Figure 6 demonstrates all possible sub-categorizations of the category $C(x^i) = C(u^p v^q)$: line 1 ($p = 0$, $q = 3$) and its fan of segments display $2^p = 1$ of subcategories $C_{00}(v^3)$ of $2^q = 8$ of items $(c_{u0}u_0^0)v_s^3 = c_{x3}x_j^3$ of definite meanings; line 2 ($p = 1$, $q = 2$) and its fan of segments show $2^p = 2$ of subcategories $C_{1r}(v^2)$ of $2^q = 4$ of items $(c_{u1}u_0^1)v_s^2$ of definite meanings and the only subcategory $C_2(u^1)$ of $2^p = 2$ of synonyms $c_{u1}u_r^1$ of conditional meanings of 2-level meaning uncertainty; line 3 ($p = 2$, $q = 1$) and its fan of segments depict $2^p = 4$ of subcategories $C_{2r}(v^1)$ of $2^q = 2$ of items $(c_{u2}u_0^2)v_s^1$ of definite meanings and the only subcategory $C_1(u^2)$ of $2^p = 4$ of synonyms $c_{u2}u_r^2$ of conditional meanings of 1-level meaning uncertainty. For the category $C(x^i) = C_{00}(v^i)$ and for all its subcategories $C_q(u^p)$ and $C_{pr}(v^q)$, the number of their zero-level words having definite meanings, regardless of the sort of their particular sub-categorization, always remains the same because $2^i = 2^{p+q}$ (in Figure 6, $2^i = 8$).

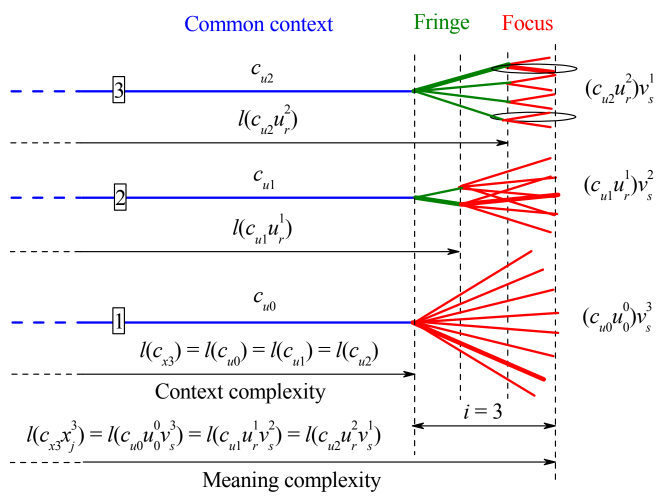

Figure 6. A category $C(x^i) = C(u^p v^q)$ of meaningful words and subcategories of them $C_q(u^p)$ and $C_{pr}(v^q)$, the case of $i = 3$. The largest common context, fringe and focal words are respectively shown as blue, olive and red line segments; thick line segments coincide completely bit-by-bit; definite meanings of circled words are directly incomparable, the distance between neighbor vertical dashed lines equals 1 bit. Other designations are as in Figure 3. See text for details.

Since synonyms are always defined given an infinite context, they should also be compared given the context. For example, in Figure 6, for subcategories $C_{00}(v^3)$, $C_{1r}(v^2)$ and $C_{2r}(v^1)$, their items shown as thick lines are bit-by-bit equivalent and the meanings of focal words $v_s^3 = x_j^3$, v_s^2, and v_s^1 of these strings may directly be compared. Since these focal words have different fringes or immediate contexts u_0^0,

u_r^2, u_r^1 and, consequently, different resulting total contexts $c_{u0}u_0^0$, $c_{u1}u_r^1$, $c_{u2}u_r^2$, they have different definite meanings and name the same thing's features (parts, attributes, properties, traits) under condition that "visual area" is gradually shrinking and covers gradually smaller number of "visible" features (from 8 for v_s^3 to 2 for v_s^1). Words that are circled in Figure 6 have different contexts and as a result directly incomparable meanings. Such a comparing should conditionally be performed.

2) Ontologies/conceptual spaces of meaningful words and hierarchies of brain devices for processing these words: We refer to particular arrangement of synonyms of definite and conditional meanings that are the members of all focal, $C_{pr}(v^q)$, and fringe, $C_q(u^p)$, subcategories of particular category, $C(x^i) = C(u^pv^q)$, as particular BSDT PL *partial* or *sub-ontology of meaningful words* that name the features of particular multi-feature thing of the world (in Figure 6 lines 1, 2, and 3 with their fans show three possible for this example sub-ontologies). Taken together sub-ontologies or conceptual subspaces devoted to meaningful naming the features of a particular multi-feature thing produce *the whole ontology or conceptual space* of words specifying this thing (all fans taken together in Figure 6). As has been mentioned earlier, in humans, the biologically plausible number of components of composite words is expected to be 3 to 5 [44] and, consequently, biologically plausible BSDT PL ontology is expected to contain 3 to 5 levels of subcategories. Zero-level subcategories represent the synonyms of definite meanings, all other the ontology's subcategories consist of items of conditional meanings with the level of meaning uncertainty q that has 3 to 5 grades of values. The case of more than two constituents of composite words as well the case of multiple multi-feature things will not be discussed.

Each synonym related to a given sub-ontology or ontology is processed in an animal's brain by a synonym-specific network/ASM device devoted to processing this synonym only. We refer to the arrangement of such real-brain recognition devices serving the sub-ontology or ontology as a real-brain network/ASM *sub-hierarchy* or *the whole hierarchy* (neural subspace or the whole neural space [28]) devoted to recognize and process the patterns of signals originated from different features of a multi-feature thing whose sub-ontology or ontology of names implements this network/ASM sub-hierarchy or hierarchy. Regardless of whether they are in focal or fringe positions, between the ontology's separate words (patterns of +1s and –1s), the hierarchy's separate recognition devices, and the separate features of a multi-feature thing of the world a one-to-one correspondence exists. It is assumed the hierarchy's recognition devices are real-brain implementations of BSDT ASMs (cf. Figure 8). The BSDT even makes a distinction between the ASMs dedicated to the recognition of fringe and focal constituents of composite words; they are *passive ASMs* and *active ASMs*, respectively [30]. As usually, the ontology's zero-level (focal) constituents have definite meanings (they are processed by active ASMs), all the other its constituents have conditional meanings and are processed by passive ASMs. This explains why we effortlessly recognize and name the whole multi-feature things (e.g., human faces) or their separate salient features (e.g., eyes or lips) and why we experience difficulties when recognizing and naming the relationships between these features or between these features and the whole thing. The same concerns the arguments for their allocation.

Multi-feature-thing-specific ontologies and network/ ASM hierarchies may dynamically be constructed for temporal purposes in the process of an animal's adaptation to permanently changing environment (e.g., a short-term memory for the traffic on a street cross) and may almost "for ever" be embedded ("hardwired") into an animal's anatomy in the process of animal evolution, development and, finally, learning from experience (e.g., long-term memory for faces).

3) Semantic rule of identity: Given the context, the category/subcategory's synonyms name different sub-devices of the same multi-purpose real-brain recognition device and simultaneously respective features of the multi-feature thing of the world generated the signals this device is devoted to process. For this reason, direct *interchangeability* of synonyms is only possible among the members of the same category/subcategory (Figure 6): the change of a synonym changes the choice of an animal's focus of attention (Sections V B2 and VI B2) and changes the feature of current interest of *the given* multi-feature thing of the world – that is the BSDT PL's semantic *rule of identity*. In the BSDT PL, there is in principle no possibility of ascribing different names to the same thing or to the same feature of this thing because by definition each meaningful word is unique and its meaning is actually keeping in the mind of behaving organism. If it is not the case a malfunctioning of the organism appears.

4) Randomness and irreducibility of synonyms: To name a feature of a thing by one of 2^i of the category's synonyms x_j^i from the set S_{xi} (to teach one of 2^i of the hierarchy's sub-hierarchies to recognize the x_j^i), particular x_j^i (particular network/ASM sub-hierarchy devoted to store and recognize an i-bit message) is chosen *in random* because none of the category's features to be named, none of binary patterns x_j^i that may be used for their naming, and none of the network/ ASM sub-hierarchies that may be chosen and taught to recognize the x_j^i have any priority over the others. Hence, for naming the features of a multi-feature thing, BSDT PL category's synonyms can only be chosen from their given range of values *in random*. But, on the other hand, once random choice of a name for naming a thing has been done (particular network/ASM sub-hierarchy has been taught to store and recognize a name), this name (respective sub-hierarchy) is then *rigidly* associated with one particular thing of the world or the thing's attribute. Thanks to this BSDT PL's peculiarity, it reconciles the notions of rigidity and contingency of names usually treated in semantics as different, e.g., [48].

The category's synonyms x_j^i can also be understood as *natural numbers* ranged from zero to $|S_{xi}| - 1 = 2^i - 1$ or from $2^i - 1$ to $2(2^i - 1)$ and written in binary string notations (see (1) in Section VII A). Since the synonyms constituting a category/subcategory are *randomly chosen natural* *numbers*, they cannot be reduced to simpler mathematical expressions and are consequently *irreducible* (cf. Figure 8).

VI. BSDT PL Real Numbers, Continuum, and Continuity-discreteness Unity and Uncertainty

BSDT PL meaningful words could further be interpreted as traditional mathematical structures but written in a non-traditional way. This entails essential consequences.

A. Real Numbers and the Cantor's Continuum Hypothesis

BSDT PL meaningful words are traditional real numbers but have a common infinite beginning.

1) Real numbers and their countable continuum: All the strings $c_{xi}x_j^i$, generated by the same master string c_{x0}, have a common infinite beginning of the same infinite length $\aleph_0$ and taken together they produce the proper class S_{cx0}, $c_{xi}x_j^i \in S_{cx0}$. The number of elements of the *i*th fraction of the S_{cx0} (it comprises all the $c_{xi}x_j^i$ with affixes x_j^i that are not longer than *i* bits) is given by the sum $|S_{cx0}^i| = \sum 2^k = 2^{i+1} - 1$ where $|S_{cx0}^i|$ is the cardinality of the fraction S_{cx0}^i and $k = 0, 1, \ldots, i$. Consequently, between natural numbers in their usual order and all the members of the S_{cx0} a one-to-one correspondence can be established (see also (1) and Section VII A). That means the cardinality of the S_{cx0}, $|S_{cx0}|$, and the cardinality of the totality of natural numbers, $\aleph_0$, are equal to each other, i.e., $|S_{cx0}| = \aleph_0$. On the other hand, if to posit $i = \aleph_0$ and neglect the 1s in above expression for the $|S_{cx0}^i|$ (they are in this case inessential) then it gives the cardinality of the S_{cx0} as $2^{\aleph_0}$, i.e., $|S_{cx0}| = 2^{\aleph_0}$ where $2^{\aleph_0}$ is the size of the Cantor's continuum.

The members of the proper class S_{cx0} do not exhaust the totality of all the BSDT PL's possible meaningful strings. Along with the master string c_{x0}, any of its infinite fractions c_{xI} that shares with the c_{x0} its infinite beginning but is shorter in *I* bits also generates its own proper class S_{cxI} of the size $\aleph_0 = 2^{\aleph_0}$ (I = 1, 2, 3 and so on without an upper limit). Each proper class S_{cxI} with $I > 0$ has the same number of items as the S_{cx0} (S_{cxI} with $I = 0$) but comprises meaningful words that have *I*-bits-smaller meaning complexity with respect to meaningful strings of the S_{cx0}. Hence, the totality of BSDT PL meaningful strings consists of all the members of all proper classes S_{cxI} that produce together *BSDT PL continuum – the totality of all its real numbers* (to remind, BSDT PL meaningful binary strings $c_{xi}x_j^i$ of the length of $\aleph_0$ bits can be understood as real numbers written as one-way infinite strings with common infinite beginnings and *i*-bits-in-length explicitly specified right-most fractions). As the number of elements of each of the classes S_{cxI} with $I = 0, 1, 2, \ldots$ is $\aleph_0 = 2^{\aleph_0}$ and the number of such classes is $\aleph_0$, the size of the totality of BSDT PL real numbers is also $\aleph_0 = 2^{\aleph_0}$. In other words, *BSDT PL continuum is countable*.

2) Refutation of Cantor's continuum hypothesis: The countability of the BSDT PL continuum radically contradicts to Cantor's continuum theory stated the existence of *two* kinds of infinities: the countable infinity of natural numbers (the cardinality of their totality is $\aleph_0$) and the uncountable infinity of real numbers (the cardinality of their totality is $2^{\aleph_0}$). With the help of its diagonal argument Cantor found that $\aleph_0 < 2^{\aleph_0}$. Additionally he conjectured his *continuum hypothesis*, *CH*, stating that there is no infinite set whose cardinality would be in between $\aleph_0$ and $2^{\aleph_0}$. But thanks to Kurt Gödel [49] and Paul Cohen [50], [51], it is known the CH is independent on the ZF or ZFC and can be neither proved nor disproved assuming the ZF/ZFC holds. It means simultaneously such extensions of ZF/ZFC are possible for which the CH either holds or fails. BSDT PL extension of the ZF/ZFC by the hypothesis of concurrent infinity is in this respect special because it leads to the *countable* continuum. This property of the continuum contradicts to Cantor's diagonal argument but is highly desirable of the view of such mathematicians as, e.g., Leopold Kronecker, Henri Poincaré, L. E. J. Brouwer or Hermann Weyl who were the opponents of Cantor's continuity/infinity theory. Thus, BSDT PL rather refutes than solves the CH or the first David Hilbert's problem [52].

Why in the case of the BSDT PL Cantor's diagonal argument does not work can be explained in the following way. We assume, as Cantor did, all the real numbers preexist and may be presented as infinite symbolic strings of the length $\aleph_0$. Cantor also postulated that each symbol in these strings is known and at any moment its identity, if one desires, may immediately be disclosed at least in principle. But contrary to Cantor, for each real number, we assume that the amount of its string symbols whose identities are known and at any moment may immediately be disclosed is always finite, whereas the amount of symbols of unknown identity is infinite. The reason for this fundamental distinction is that, in string descriptions of BSDT PL real numbers, the amount of their explicitly known symbols is always *finite* because they are defined by an always *finite* process of their construction.

B. Continuity-discreteness and other Related Unities and Uncertainties, Elusiveness of the Focus of Attention, the Symbolism's Insufficiency

Meaningful words in the form of real numbers sharing their infinite beginning provide a new view of possible symbolic representations of knowledge and computations.

1) Continuity-discreteness, quality-quantity, and reality-symbolism unities and uncertainties: The constructability of BSDT PL real numbers and the countability of the totality of them do not exert any influence on practical computations because they are in fact always performed with infinite real numbers presented by finite binary strings only. But the very fact of knowing the real numbers' constructability is of great practical importance because it shows that though the amount of symbolic information about a thing is always *infinite* its *finite* part may only be known at any moment.

A particular meaningful pattern of *i* signals originated from a thing of the world is processed by an animal's recognition device (an implementation of the BSDT ASM [30]) deliberately designed for this aim. The complete description of this device and the meaning of the pattern of signals it recognizes are given by an infinite binary string $c_{xi}x_j^i$ whose finite *i*-bits-in-length fraction x_j^i represents the pattern to be recognized. Consequently, *the finite* fraction of the $c_{xi}x_j^i$, x_j^i, does represent currently relevant symbolic, explicitly formalized, quantitative, discrete-valued information about the thing. The remaining *infinite* fraction of the $c_{xi}x_j^i$ and *infinite* amount of symbolic information it contains, c_{xi}, are presented in a non-symbolic, informal, qualitative, implicit, real-valued, continuous way as the mentioned real-world recognition device. Hence, complete representation of any finite symbolic *meaningful* message should always consist of this message itself and of the real-world device devoted to recognizing it (cf. Section X B and D). The reason why we invoke such *unity or complementation of symbols and things* (in other cases the items of incommensurable domains) is our infinity hypothesis and our phenomenology equating infinite strings of symbols and real-world things (boxes 1 and 4, Figure 2).

Due to inherent unity of BSDT PL domains of symbols and of real-world things/real-brain devices, any *meaningful* discrete-valued description of any thing should inevitably contain some traces of continuity caused by intrinsic connections between meaningful patterns of symbols (finite vectors x_j^i) and real-brain things/devices (infinite strings $c_{xi}x_j^i$) devoted to process them. Purely discrete-valued symbolic messages are always meaningless, as in the case of Shannon [45]. BSDT PL messages or natural numbers or finite symbolic strings are meaningful because they are always conditioned by real numbers or one-way infinite strings of symbols. Of this follows a *continuity-discreteness* or *quality-quantity* or *reality-symbolism unity and uncertainty* of BSDT PL meaningful finite-in-length symbolic messages (cf. Section X E). That is why the world's BSDT PL discrete (quantitative, symbolic) and continuous (qualitative, real-world) representations must compete and coexist at anytime and anywhere. It is indeed the fact in real brains where cooperative and simultaneously competitive discrete-continuity effects are ubiquitous (it is sufficed to recall, e.g., close relationships between spike and wave neuron activities [53], [54]). Research concerning brain mechanisms and brain organization of uncertainty are reviewed, e.g., in [55]. An example is in Section X E.

2) Elusiveness of the focus of attention: In the domain of symbols, information exchange is performed by finite symbolic meaningless messages to be processed by Turing methods. In the domain of real-world devices, communication is performed by unspecified (e.g., chemical [21]) non-symbolic messages that can not be processed by Turing methods. Isolated symbolic messages take their meanings from their interactions with the domain of real-world devices. Such interactions define for an animal/ human the sort and the amount of currently relevant meaningful symbolic information to be communicated. Simultaneously, they define the sort and amount of non-symbolic but potentially symbolic contextual information that is not communicated together with symbols but is crucial for giving them meanings. The mechanism of selecting the relevant symbolic information (e.g., right-most line segments designated the whole words x_j^i or their "focal" fractions v_s^q in Figures 3 to 6) resides in the domain of things and defines for a perceiving agent/living organism the choice of its current *focus of attention*. Recent studies indeed demonstrate the importance of non-symbolic, e.g., wave-like interactions observed by EEG (electroencephalogram) or/and fMRI (functional magnetic resonance imaging) methods that are essentially continuous and implement large-scale network interactions supporting attention mechanisms in humans, e.g., [56], [57]. Moreover, a current empirical finding seems to directly indicate [58] that the purely symbolic approach, restricted to considering the spike brain activity only, is insufficient to explain the effects of attention and, consequently, "other processes must have a key role" [58], [59]. The reason is that an "SC inactivation caused major deficits in visual attention tasks" while simultaneous "attention-related effects in MT and MST remain intact," despite the usual view of selective spike activity of neurons in MT and MST as the main correlate and distinctive feature of visual attention [58]. SC, MT, and MST are respectively the superior collicus, middle temporal, and medial superior temporal monkey brain areas involved in motion-detection tasks.

3) Incompleteness of the BSDT PL formalism: BSDT PL discrete-valued formalism informs nothing of mechanisms of the arrangement of its composite words or sentences (Section V B2) and of selecting their fragments that are to be placed into the current focus of an animal's attention (Figure 4). Thus, in spite of its perfection and efficacy (Sections X and XI), discrete part of BSDT PL formalism is *incomplete* and, consequently, *insufficient* to ensure its own running in full. Its incompleteness is a manifestation of the continuity-discreteness or symbolism-reality uncertainty predicted by the BSDT PL. It is this uncertainty that is the reason why complete *symbolic* theory of anything, including the BSDT PL itself, is fundamentally impossible.

VII. BSDT PL ARITHMETIZATION BY NATURAL NUMBERS AND ITS ESSENTIAL RANDOMNESS

Considering the right-most finite fraction of meaningful words as natural numbers provides unexpected solutions to some mathematical problems of great generality.

A. Arithmetization by Natural Numbers of Mathematical Expressions of Different Meaning Complexity

Every $c_{xi}x_j^i \in S_{cx0}$ generated by master string c_{x0} is uniquely labeled by its affix x_j^i or by its indices i and j. The x_j^i encodes i and j as its own length and its own *content* understood as an arrangement of this vector's positive and negative components. As it is usually done in computer sciences (see item two in (1)), strings x_j^i may also be treated as natural numbers written in binary notations and ranged from zero to $2^i - 1$. In such a form, the x_j^i with different i but same j correspond to same natural numbers. To ensure the unique bijection from strings to numbers, let us introduce decimal equivalents to strings x_j^i, the numbers G_{ij}^{x0}, as

$$G_{ij}^{x0} = \sum_{k=1}^{i} 2^{k-1} + \sum_{k=0}^{i-1} (x_j^i(k)+1)2^{k-1} \qquad (1)$$

where $x_j^i(k)$ is the kth component of the x_j^i (it equals either +1 or –1); the second item of the sum (1) is, for this x_j^i, the value of j given the value of i. At $i \geq 1$, G_{ij}^{x0} = 1, 2, 3 and so on without an upper limit.

The affixes x_j^i of all the conceivable strings $c_{xi}x_j^i \in S_{cx0}$ may be treated as all the conceivable written in binary notations meaningful (given the c_{xi}) or meaningless (if the c_{xi} is ignored) mathematical expressions/assertions not longer than i bits. It means, strings x_j^i provide the labeling of themselves and of the mentioned expressions and consequently may be considered as *Gödel vectors/strings* that, because of (1), are one-to-one related to *BSDT PL Gödel numbers* G_{ij}^{x0}. Gödel numbers (1) are natural numbers and enumerate themselves and all the conceivable mathematical meaningful expressions $c_{xi}x_j^i$ or meaningless expressions x_j^i. Consequently, the numbers G_{ij}^{x0} provide for these expressions their complete *arithmetization by natural numbers*.

The G_{ij}^{x0} enumerate all the members of a given proper class, S_{cx0}. But the totality of such classes generated by different master strings c_{xI} is infinite (Section VI A) and for each of them, S_{cxI}, its own system of Gödel numbers G_{ij}^{xI} can by analogy be defined ($I = 0, 1, 2, ...$; if $I = 0$, $G_{ij}^{xI} = G_{ij}^{x0}$). At different values of I, the totalities of strings x_j^i and natural numbers G_{ij}^{xI} are the same but they enumerate meaningful mathematical expressions of different meaning complexities (Section V C). The arithmetization just introduced is a non-Gödelian one though already in 1946 Kurt Gödel seemed to have envisaged something similar when he said about the possibility "to take the ordinals themselves as primitive terms" [60].

B. Natural Numbers, Gödel Numbers, Omega Numbers and Essential Randomness of their Use as Names

Every $c_{xi}x_j^i \in S_{cx0}$ contains an infinite-on-a-semi-axis and *always unspecified* initial part c_{xi} and a finite fraction x_j^i that, given the value of i, has *random* (incompressible and incomputable, Section V D4) arrangement of its binary components. This property of BSDT PL meaningful strings or *real numbers* (Section VI A) reflects their constructability ("random computable enumerability" [61]) and, given the value of i, the randomness of the x_j^i (its "algorithmic randomness" [62]). The x_j^i is to be processed by a special-purpose *self-delimiting computer* existing, we suppose, as a BSDT ASM [30] and dealing with random binary computer algorithms not longer than i bits. This property of BSDT PL words reflects their inherent connections with the devices that process them in the best way, of course, if their previous learning was perfect [31]. The mentioned properties of strings $c_{xi}x_j^i$ demonstrate that they are in fact binary string representations of *computably enumerable random real numbers* that, according to [61], are simultaneously *Omega-like* and *Omega numbers*. Gregory Chaitin's Omega number Ω gives the halting probability of randomly chosen binary algorithms running on a computer given its hardware and software [62]. The affix of the $c_{xi}x_j^i$, x_j^i, is an i-length fraction of Ω or a "partial" Ω, Ω_{ij}^{x0}, providing the halting probability of random binary algorithms running on the ijth self-delimiting in i bits computer. Omega-like numbers were introduced as a generalization of Ω but it later turned out, they and another generalization of Ω known as enumerable random real numbers are equivalent to Ω numbers [61].

The ijth numerically written partial halting probability can be presented [62] as

$$\Omega_{ij}^{x0} = \sum_{k=0}^{i} (x_j^i(k)+1)/2^{k+1} \qquad (2)$$

where x_j^i is the same probability written in binary string notations, $x_j^i(k)$ is the kth component of the x_j^i (it equals either +1 or –1), k is the length of a random binary algorithm running on the ijth computer (we suppose, BSDT ASM [30]), $1/2^k$ is the probability that this algorithm halts on this computer (if it halts, $(x_j^i(k) + 1)/2 = 1$ otherwise $(x_j^i(k) + 1)/2 = 0$). Halting probabilities Ω_{ij}^{x0} correspond to $c_{xi}x_j^i$ generated by the master string c_{x0}. Master strings c_{xI} whose meaning complexities are I bits smaller than the meaning complexity of the c_{x0} generate proper classes S_{cxI} and partial halting probabilities Ω_{ij}^{xI} with $I = 0, 1, 2, \ldots$ (if $I = 0$, $S_{cxI} = S_{cx0}$ and $\Omega_{ij}^{xI} = \Omega_{ij}^{x0}$; cf. Section VI A). At different values of I the totalities of strings x_j^i and halting probabilities Ω_{ij}^{xI} are the same but concern randomly chosen meaningful algorithms of different meaning complexity.

BSDT PL words x_j^i are *simultaneously* natural numbers, Gödel numbers, random algorithms, and partial Ω written in binary notations. That is, BSDT PL Gödel numbers G_{ij}^{xI} are essentially *random-valued*. This "strange" property can be explained if one notices that given the i the values of their indices j are randomly chosen from the range of zero to $2^i – 1$ or (see (1) and Section V C4) from $2^i – 1$ to $2(2^i – 1)$ where $2^i – 1 = G_{ij}^{x0}$ at $j = 0$ and $2(2^i – 1) = G_{ij}^{x0}$ at $j = 2^i – 1$. In other words, from 2^i of equal in rights ways of the enumeration of indices j, the second item in (1) gives only the one that was *randomly* chosen here for the reason of its analytical convenience only. As the size of the totality of different x_j^i is equal to the size of the totality of natural numbers $\aleph_0$ (Section VI A), the totalities of BSDT PL Gödel numbers (1), Chaitin numbers (2), random binary algorithms and self-delimiting computers devoted to process them are also of the size $\aleph_0$. As it was first demonstrated by Alan Turing [7], halting probabilities are incomputable. According to [62], they may be treated as true, unprovable assertions or irreducible mathematical facts (axioms) that in our case represent/name the things of the world or their particular features.

The totality of meaningful strings $c_{xi}x_j^i$ may be interpreted "as an alphabet of human thought" with the help of which "everything could be described and distinguished by means of the combination of the letters of this alphabet," to use words of Gottfried Leibniz (quotations from [63, p. 56]).

VIII. BSDT PL Thruth and Understanding the Truth

Defining and confirming the truths of meaningful words is inevitably needed for their successful practical use.

A. Convention on Truth

The name x_j^i is true if the truth value $T(c_{xi}x_j^i)$ of its meaning $M(x_j^i) = c_{xi}x_j^i$ is "true" or, in other words, if strings c_{xi} and x_j^i are correctly joined to each other. If there is no such correct correspondence, the truth value $T(c_{xi}x_j^i)$ of meaningful name is "false".

B. Completeness of Truths and Gödel's Incompleteness

Since the cardinality of the S_{cx0}, $|S_{cx0}| = \aleph_0$, is infinite, the number of BSDT PL truths is also potentially infinite and, for any meaningful string, its truth value $T(c_{xi}x_j^i)$ certainly exists and equals either "true" or "false". Each true meaningful name (infinite symbolic representation of a primary thought), e.g., $c_{xi}x_j^i$ names by definition the *ij*th real-world thing given to an animal through its *ij*th psychological state (physical implementation of the primary thought) or, in other words, through the activity of physically implemented real-world BSDT ASM devoted to process the x_j^i, ASM(x_j^i) [28], [30]. Thus, for meaningful words, the truth is the norm and the falsity is an anomaly caused, e.g., by an animal's dysfunction or disease. In any case, there is *no* lie and *no* liar paradox – a source of Kurt Gödel's incompleteness [6] which does not hold for BSDT PL *meaningful* [30], [32] names $c_{xi}x_j^i$ (Section VII). Axioms, theorems, and metamathematical expressions/assertions of a formal axiomatic system for which the Gödel's incompleteness holds are in our terms an infinite fraction of infinite in number *meaningless* strings x_j^i [17], [32]. These inferences are caused by the fact that BSDT PL name meanings are always the ones that animals/humans keep *actually* in their mind, like meanings of hypothesized by W. V. Quine "eternal" sentences [64]. It means, to survive, an animal does not lie to itself and it is the reason why the BSDT PL works so well as a primary language or a language of primary thoughts [28], [65], [66]. At the same time, a zero-level name's fringe words, due to their colored non-locality, have no definite but *conditional* meanings (Section IX and Figure 7). Fringe words are fuzzy/vague in meaning or "in limbo" in words of Quine [64], but their truths are *not* conditioned and always remain of certain values. The vagueness of meanings of fringe names is a BSDT PL manifestation, for the case of infinite sequences, of the famous *Burali-Forti paradox* (Section IX C) but it does not concern the truths.

BSDT PL convention on truth essentially differs from Alfred Tarski's *convention T* [67]. Tarski's definition is in fact *syntactical* and holds for an axiomatically defined pair object-language/meta-language only, whereas BSDT PL definition is *semantical* and uses the real world for checking the truths. Truth values of BSDT PL names are *unique* and *conclusive*. Any hierarchy of these truths is neither possible nor required, contrary to Tarski's syntactical approach implying that for any meta-language its meta-meta-language can in turn be conceived and so up. For this reason, for each of Tarski's meta-languages its higher-level meta-meta-language and respective higher-level truth (the truth of sentences of respective higher-level meta-language) could in general always be defined. Here, in words of Quine, "there is interlocking of class hierarchy with truth hierarchy" [64, p. 90], which may be traced back to early Bertrand Russell's theory of types [68]. Hence, Tarski's truths are *relative* while BSDT PL truths are *absolute*.

C. Discovering, Understanding and Confirming the Truths

The BSDT PL truths are here introduced as a *correspondence* between names (finite strings x_j^i) and things of the world (infinite strings c_{xi} or $c_{xi}x_j^i$). Indeed, each true meaningful name $c_{xi}x_j^i \in S_{cx0}$ names by definition the *ij*th real-world thing given to an animal through its *ij*th psychological state or the activity in the *ij*th BSDT ASM, ASM(x_j^i), designed, implemented in a physical form, and learned beforehand to recognize/select exactly the *ij*th thing by its symbolic name x_j^i [1], [17], [32]. As truth value $T(c_{xi}x_j^i)$ is never communicated together with the x_j^i, it should always be discovered in the process of *decoding* or *understanding* the meaning of the received name x_j^i and confirmed by checking the correspondence of this name to that reality or, more accurately, to an animal's psychological state (an activity of respective recognition device) represented this reality. In living organisms, it is most probably done by physical/anatomical segregation and specification of communication channels (input/output sensory submodalities) or/and by the choice of different physical carriers for different types of symbolic signals to be communicated. By means of such segregation and specification, the required *ij*th neural subspace/ASM hierarchy (the *ij*th computer for particular mental computations, Section VII B) is eventually allocated. By the following convergence of relevant channel-specific symbolic information from different communication channels an integral or holistic and, consequently, most reliable estimation of the current state of the animal's internal or/and external environment has to be achieved.

It is supposed the hierarchies (ontologies) of meaningful names/strings but *not* their truth values are implemented in the brain by means of BSDT neural subspaces/ASM hierarchies for signal processing, memory, decision-making and consciousness [28], [31], [32]. The truth value of each meaningful word of such hierarchy is not a property of the organism's device serving this word but a result of *evaluating* the state of this device. For this reason, of *the third person perspective,* it exists as a psychological state of an external observer who should intentionally define/ discover this value of truth ("true" or "false") by comparing the result of running the device serving the word of interest and respective thing of the world. Of *the first person perspective*, each meaningful word's truth value is postulated to be "true" because in terms of truths this fact reflects simply a distinctive feature of the definition of such words, namely that the word's meaning is the one that an animal actually stores in its mind.

IX. BSDT PL Meaning Ambiguity

Meaningful words are always the members of a proper class and this exerts essential influence on the possibility of comparing their meanings. In particular, in many cases, meanings are directly incomparable and, consequently, meaning ambiguities are inevitable.

A. Definite Meanings of Simple Words

It is assumed that, in a meaningful string $c_{xi}x_j^i$, its context c_{xi} and its simple focal name x_j^i describe respectively the *static* part or a "hardware" of the ASM(x_j^i) already fixed in the course of evolution and its *dynamic* part or "software" designed in the course of the hardware's adaptive learning and development. The length of x_j^i in bits, i, defines the number of now essential (explicitly considered) features of the ijth thing named by the x_j^i; the jth arrangement of ±1 components of x_j^i is the jth BSDT PL description of this ijth thing (e.g., the value +1 or –1 of a component of the x_j^i may mean that the respective feature is included to, +1, or excluded from, –1, the consideration). The complexity of meaning of the name x_j^i reflects the meaning complexity of the physically implemented real-world ASM(x_j^i) and an organism of which the ASM(x_j^i) is a part but not the complexity of the thing named by x_j^i.

B. Definite and Conditional Meanings of Constituents of Composite Words

Different constituents of composite words are the members of different proper classes or of different meaning complexity. For this reason, relations between their meanings may be rather intricate.

1) "Virtual" devices for processing the "virtual" things: If the string x_j^i is a composite one, $x_j^i = u_r^p v_s^q$, then strings $c_{up}u_r^p$ and $(c_{up}u_r^p)v_s^q = c_{vq}v_s^q$ describe, given the context $c_{xi} = c_{up}$, an ASM(u_r^p) and ASM(v_s^q) that may for a time period dynamically be created from the ASM(x_j^i) that in turn is the product of a similar process described by the string $c_{xi}x_j^i$. ASM(u_r^p) and ASM(v_s^q) are "virtual" ASMs (i.e., temporally designed for) selecting the names u_r^p and v_s^q of the prth and the qsth "virtual" things (i.e., of temporally highlighted/ allocated fractions of the ijth composite thing named by its ijth composite name x_j^i). In other words, virtual ASMs highlight the prth and qsth "partial" meaningful fractions of the ijth description of the ijth thing (cf. Figure 6). Composite names essentially enrich the BSDT PL semantics but raise the problem of comparing the meanings of names selected by ASM(u_r^p), ASM(v_s^q), and ASM(x_j^i).

2) Comparing the meanings of whole composite words and their focal fractions: Zero-level names x_j^i and v_s^q (v_s^q is a part of the $x_j^i = u_r^p v_s^q$) name given the context *the same* thing in the same way but from different points of view defined by their contexts (static for x_j^i, c_{xi}, and in part dynamically created for v_s^q, $c_{vq} = c_{up}u_r^p$; Figure 7(a)). The v_s^q is selected under condition $c_{xi} = c_{up}$ (for x_j^i and v_s^q their common context is c_{xi}) by the ASM(v_s^q) that is "virtual" with respect to the ASM(x_j^i). Thus, the ASM(x_j^i) can temporally serve as the ASM(v_s^q) but in any case the same thing is under the consideration and the meaning of x_j^i, $M(x_j^i) = c_{xi}x_j^i$, and the meaning of v_s^q, $M(v_s^q) = (c_{up}u_r^p)v_s^q = c_{vq}v_s^q$, may unambiguously be related. As thick line segments in Figure 6 demonstrate, a $c_{vq}v_s^q$ is simply another realization of the $c_{xi}x_j^i$.

3) Comparing the meanings of whole composite words and their fringe fractions: If u_r^p is a q-level fringe of zero-level focal string v_s^q and they are the fractions of the $x_j^i = u_r^p v_s^q$ (Figure 7(a)) then u_r^p has *no* definite meaning (Section V B and C). But it could get a *conditional meaning* if one supposes that u_r^p is conditioned by the color of a colored zero-level name $u_r^p(color)$ selected by a respective q-stages-back-in-evolution ASM. If it is, uncolored zero-level names x_j^i in Figure 7(a) are unambiguously related to colored zero-level names $u_r^p(color)$ in Figure 7(b). Vectors $u_r^p(color)$ and u_r^p conditioned by one of the q colors *color* have conditional but certain meanings. But once colors are deleted (only uncolored strings are used in computations) the one-to-one correspondence between x_j^i and $u_r^p(color)$ disappears and, instead of it, we obtain 2^q-state uncertainty between the x_j^i and u_r^p and between the definite meaning of x_j^i and conditional meaning of u_r^p (Figure 7).

4) Comparing the meanings of words naming evolutionary predecessors and successors: If, given the context $c_{xi} = c_{up}$, names x_j^i and u_r^p (or y_r^p in Figure 5) are both of the level of zero, then their meanings are to be of *different* proper classes and should have *different* meaning complexities (to remind, meaning complexity of x_j^i is $l(c_{xi}x_j^i) - l(c_{up}u_r^p) = i - p = q$ bits larger than that of u_r^p; see Figures 4 and 5, Figure 7(a) and (c)). This means they describe *different* things from the same point of view or the same thing at *different* stages of its evolution. The names x_j^i (Figure 7(a)) and u_r^p (Figure 7(c)) are respectively selected by present-stage-of-evolution ASM(x_j^i) and q-stages-back-in-evolution ASM(u_r^p) and refer to animals of evolutionary different species. Meaningful string $c_{up}u_r^p$ and respective part of $c_{xi}x_j^i = (c_{up}u_r^p)v_s^q$ may coincide bit by bit but even in this case meanings of x_j^i and u_r^p may only *conditionally* be related to each other and 2^q additional conditions (strings v_s^q in Figure 7(a)) are required to uniquely establish their correspondence.

5) Graphical illustration of meaning ambiguities:

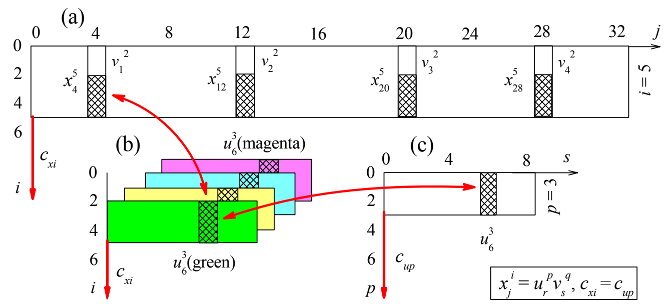

Figure 7. Comparing given the context different-level BSDT PL meaningful names of different proper classes: (a) zero-level names, (b) colored zero-level names corresponding to names in (a), (c) zero-level names that are predecessors to names in (a) and counterparts to names in (b).

In Figure 7, panel (a) demonstrates zero-level names x_j^i and v_s^q of meaningful strings $c_{xi}x_j^i$ and $(c_{xi}u_r^p)v_s^q$ ($x_j^i = u_r^p v_s^q$; $i = 5$, $p = 3$, and $q = i - p = 2$); the rectangle has the height i bits and the width $|S_{xi}| = 2^i = 32$ bits, the ijth bar of the height i in the jth horizontal position designates the name x_j^i that in a numerical form (see (1) and (2)) corresponds to Gödel number G_{ij}^{x0} and partial Chaitin number Ω_{ij}^{x0}; bars $x_4^5 = u_6^3v_1^2$, $x_{12}^5 = u_6^3v_2^2$, $x_{20}^5 = u_6^3v_3^2$ and $x_{28}^5 = u_6^3v_4^2$ that correspond to four colored highlighted bars in (b) are also highlighted (u_6^3 is q-level fringe of zero-level names v_s^q that is a focal fraction of the x_j^i); substrings v_1^2, v_2^2, v_3^2, and v_4^2 may encode the colors of colored strings $u_r^p(color)$ in (b). Panel (b) shows conditioned zero-level names $u_r^p(color)$ that in a numerical form (see (1) and (2)) correspond to colored Gödel numbers $G_{pr}^{u0}(color)$ and colored partial Chaitin numbers $\Omega_{pr}^{u0}(color)$; under condition that the word *color* is a parameter, names $u_r^p(color)$ are uniquely related to names x_j^i in (a) and selected by conditional q-stages-back-in-evolution ASMs; colored words $u_r^p(color)$ conditionally name the things unconditionally named by the x_j^i; equal-in-size rectangles colored in $|S_{vq}| = 2^q = 4$ colors consist of $|S_{up}| = 2^p = 8$ bars of the height p; uncolored bars in (a) and respective colored bars in (b) (e.g., $u_6^3(green)$ and x_4^5) denote different descriptions of the same thing. Panel (c) displays uncolored zero-level or focal names u_r^p (y_r^p in Figure 5) of meaningful strings $c_{up}u_r^p$ that name evolutionary predecessors of things named by the words x_j^i; the prth bar of the height p in the rth horizontal position (it is shaded) designates u_r^p for the case $u_r^p = u_6^3$ (in a numerical form it corresponds to Gödel number $G_{3,6}^{u0}$ and Chaitin number $\Omega_{3,6}^{u0}$).

In (a), (b), and (c), strings that are bit-by-bit equivalent to the u_6^3 are shaded in the same way. Contexts are shown as thick arrows and equal to each other bit by bit, $c_{xi} = c_{up}$. Uncolored and colored names name real-world unconditional and real-world conditional ("virtual") things, respectively. Between names in (a) and in (b) a bijection $x_j^i \leftrightarrow u_r^p(color)$ exists that may be for example $x_{28}^5 \leftrightarrow u_6^3(magenta)$ or $x_4^5 \leftrightarrow u_6^3(green)$. A bijection also exists from names u_r^p in (c) to given-color names $u_r^p(color)$ in (b). For example, it may be $u_r^p \leftrightarrow u_r^p(green)$. But if such a bijection was already established then other conceivable bijections, e.g., $u_r^p \leftrightarrow u_r^p(magenta)$ become impossible. Once colors are deleted, these bijections (they are indicated as curved bidirectional arrows) disappear producing, instead of 2^q-state (4-state in (b)) discrete colored non-locality of vectors u_r^p, 2^q-state (4-state in (b)) uncertainty (degeneracy) of meaning relations between names in (a) and (b), in (b) and (c), and in (a) and (c).

C. *Relationships between Meaning Ambiguities and Burali-Forti Paradox*

The origin of conditional relationships between meanings of names of different meaning complexities is the properties of ultimate/proper classes caused in turn by BSDT PL infinity hypothesis or vice versa, as the hypothesis of concurrent infinity was introduced when proper classes were already known in literature, e.g., [39]. Of this follows the famous Burali-Forti paradox according to which "there can be two transfinite (ordinal) numbers, *a* and *b*, such that *a* neither equal to, greater than, nor smaller than *b*" [8, p. 157] means in our terms that meanings of BSDT PL names whose meaning complexities differ in *q* bits can only be compared with 2^q-state uncertainty. In Figure 7 infinite strings $c_{xi}x_j^i$ and $c_{up}u_r^p$ are like Burali-Forti's transfinite ordinals *a* and *b* mentioned above.

The Burali-Forti paradox reflects the meaning-ambiguity properties of BSDT PL infinite symbolic statements/strings of different meaning complexities but in terms of transfinite ordinals. On the other hand, BSDT PL provides specific quantification of the ambiguities stated for different transfinite ordinals by the Furali-Forti paradox but in terms of BSDT PL strings of different meaning complexities. The reason is in end our infinity hypothesis.

X. Numerical and Empirical BSDT PL Validation

Now it is time to consider the BSDT PL validation.

A. *Disappearing the Bounds between Mathematics and Reality*

On the one hand, given the infinitely long context or "boundary conditions", the BSDT PL performs traditional mathematical computations with finite binary messages and is certainly a kind of mathematics that we call the mathematics of meaningful computations (Section IV). On the other hand, the BSDT PL is a kind of natural science because infinite-in-length boundary conditions used in its computations are implemented as real-world physical devices and, consequently, the computations themselves contain inevitably indispensable, inseparable from the symbolism elements of reality. Resulting *symbolism-reality unity and uncertainty/dichotomy* (cf. Section VI B) is not a failure or misunderstanding, it is the inherent property and distinctive feature of the BSDT PL caused directly by the hypothesis of concurrent infinity and its phenomenology formalization. Owing to this feature, within the BSDT PL framework, the distinctions between mathematics and reality, between mathematics and natural sciences become rather vague and sometimes disappear. That is also the reason why the BSDT PL can not be validated by the traditional in pure mathematics method of formal proofs, i.e., by deriving theorems from axioms. BSDT PL computations are conditioned by infinite-in-length context and for this reason contain some elements of mind/ psychology (i.e., meanings of words) whereas standard mathematics ignores meanings by definition. That is why the only way to confirm the validity of the BSDT PL remains to compare its predictions with *real-world meaningful computations* that are abounds in living organisms. In sum, to validate the BSDT PL, it is needed to appeal to neuroscience, cognitive sciences, and psychology and compare their results with the BSDT PL predictions.

B. *Solving the Communication Paradox*

The first principal point needed to be understood is how in practice to communicate the meaningful words if they are by definition fundamentally infinite and how animals/ humans solve this problem, routinely and immediately.

1) Serving the subconscious, basic behaviors and the simplest sociality: In Section V we saw BSDT PL simple words are those BSDT PL sentences that are perceived "holophrasically" and do have definite meanings. Internal structure of such sentences (the manifold of their possible focal and fringe constituents) is ignored and, consequently, they are presented *without BSDT PL syntax*. This fact and the fact that meanings of BSDT PL names are the ones that animals/humans keep *actually* in mind [1], [17], [30], [32] make the BSDT PL an appropriate tool for the description of communication without syntax or without any language at all – the style of communication that is typical for animals and human infants, e.g., [34] and references therein.

Meanings of BSDT PL words are simultaneously animal/human primary thoughts [32], i.e., the simplest or primitive or elementary patterns of involuntary, automatic or *sub-* or *unconscious* activity of their brains. This activity is in fact the activity in a particular BSDT neural subspace/ASM subhierarchy (Sections V D and X E) that may most naturally be observed by an external observer as involuntary, automatic or *unconscious* behaviors of an animal that is under examination. We refer to these behaviors or body movements as basic or inherent ones because they truly reflect respective animal/human internal states. Among *basic behaviors* or invariant elements of a "paralanguage" [18] there are the ones (e.g., breathing or heart beating) that are truly innate and the ones (e.g., walking or directing the gaze) that originate from innate/ primitive behavioral reflexes, e.g., [69], [70], [71] but demand, after an animal's birth, for their further tuning and maturation in the course of "prepared" animal leaning and development (cf. Section XI E).

As the BSDT PL is well suited to describe not only meanings and primary thoughts but also basic behaviors, it is also capable of describing the behavioristic part [72] of an animal's cognition and based on basic animal behaviors communication without syntax or without any language at all. We hypothesize, this simplest type of communication suffices to support *the simplest* sociality that is typical in animals and human infants, e.g., [34].

2) Communication paradox: Since complete symbolic descriptions of BSDT PL meaningful names, $c_{xi}x_j^i$, are fundamentally *infinite*, during any finite time period none of them can ever be communicated in full even in principle while in fact many times a day everybody observes in others and experiences himself/herself numerous successful meaningful information exchanges. To cope with infinite symbolic messages for a finite time period, super-Turing devices with super-Turing computational capabilities are certainly required. Hence, this communication paradox [1], [17], [32] demonstrates that, in spite of the fact that real-world super-Turing computers are unknown and many experts believe even impossible [38], an everyday, routine, ubiquitous use of super-Turing computations is a norm in human meaningful communication.

3) Mirror transmitter and receiver devices for solving the communication paradox: In practice, communication paradox can be solved by appealing to BSDT infinity hypothesis (Section III and [1], [17], [32]) and the technique of BSDT ASMs [30]. On the one hand, these ASMs are devoted to process infinite-in-length meaningful messages but, on the other hand, they are special-purpose Turing computers running in the specific to each of them real-world environment. As in ASMs their programmatic and computational processes are in time completely separated, they do not waste their computational resources on serving themselves and, as any other special-purpose Turing computer, are faster than universal Turing computers [7].

But, dividing in time the programming and program running is insufficient to overcome the communication paradox. To cope with it, let us additionally suppose that in a communication process the ASM-transmitter and the ASM-receiver share in full their evolutionary history, i.e., they were designed, implemented in a physical form, and learned beforehand to perform the same meaningful function – selecting the same finite binary message x_j^i given the same infinite context or the same boundary conditions c_{xi}. If it is, and not in any other case, the meaning of x_j^i, $c_{xi}x_j^i$, is equally encoded, decoded, interpreted and *understood* by both parties and for both parties the value of its truth, $T(c_{xi}x_j^i)$, is the same. For this reason, and because the name's meaning is simultaneously a psychological state an animal experiences producing as well as perceiving this name, in the process of meaningful symbolic information exchange, the transmitter and the receiver are to be physically, structurally, and functionally *equivalent* in full or to be the "mirror" replicas or "clones" of each other (cf. Section X D). The fact that any two animal/human individuals, even identical twins or clones, always have different life-long individual experiences and, consequently, are never completely equivalent, is compensated by the tolerance of ASMs to their partial internal distortions and external noise [30], [73].

Several important BSDT PL predictions that are amenable for their empirical examination come out.

C. Coding by Synaptic Assemblies

Where meanings are essential (e.g., in living organisms) BSDT network learning paradigm "one-memory-trace-per-one-network" [28], [73] must be widespread in practice and, in particular, any memory for meaningful records must be built of the number of networks that coincides with the number of records to be stored in memory. This paradigm is not consistent with the usual desire of designers and engineers to store in a network as many traces as possible but it is the mandatory BSDT PL requirement ensuring the meaningfulness of memory records (Section V B).

A recent empirical neuroscience finding of coding by synaptic assemblies [74], [75] demonstrates this BSDT PL requirement is fulfilled in practice. In laboratory, mice were trained to perform new motor tasks. In behaving animals, changes in the number of synaptic contacts associated with learning new skills were measured. In complete accordance with the BSDT PL assumption that each new memory trace should be written down in an always new separate network or *synaptic assembly*, it turned out "that leaning new motor tasks (and acquiring new sensory experiences) is associated with the formation of new sets of persistent synaptic connections in motor (and sensory)" brain areas [76, p. 859].

D. Real-brain Mirror Neurons for Super-Turing Computations by Mirror ASMs

To ensure correct understanding of the meanings of finite symbolic messages, the ASM-transmitter and ASM-receiver that are the mirror replicas of each other are to be used (Section X B). These mirror ASMs implement *meaningful super-Turing computations*: for the transmitter and the receiver, they ensure the use of the same infinitely long "boundary conditions" c_{xi} needed to perform the previously programmed Turing computations with finite-in-length strings x_j^i, e.g., as in [28], [65], [66]. Mirror ASMs *physically* divide the infinite meaningful message to be processed into infinite, c_{xi}, and finite, x_j^i, parts and take the former into account as common for both parties "hardware", designed and *physically* implemented beforehand in the course of animal evolution and development. Thanks to this "trick", to correctly understand the meaning of the $c_{xi}x_j^i$, it is enough to correctly transmit, receive, and decode the x_j^i only. The origin and theoretical substantiation of this trick is the BSDT PL phenomenology formalization (Sections III and IV) that equates infinite symbolic strings (e.g., c_{xi} or $c_{xi}x_j^i$) with real-world physical things.

Mirror ASMs also explain why meaningful communication without syntax is successful only between animals of the same or relative species: such animals are *a priori* equipped with the same "hardware" and "software" needed to finish meaningful super-Turing computations for a finite time period (small ASM changes or distortions do not matter because of ASM tolerance to damages and noise [30], [73]). The picture described is well supported by the empirical finding and studying of *mirror neurons* – the ones that are active when an animal behaves or only observes respective behaviors of others; see, e.g., [77], [78], [79] and numerous references therein. The mirror-ASM computational system just described and the mirror-neuron circuitries already observed in animals and humans [77], [78], [79] may respectively be treated as theoretical and real-brain implementations of until now hypothetical super-Turing machines with infinite inputs [38] that are capable of computing with infinite strings, which is the same as real-valued numbers.

One would object that the scheme proposed is nothing more than a regular Turing computer because nobody saw in animals anything else than regular Turing computations. But these Turing computations are "the tip-of-the-iceberg" of genuine super-Turing computations, the overwhelming part of which remains invisible if one looks for *symbolic* computations only. In the brain, conventional free-of-meaning Turing computations are immediately transformed into meaningful super-Turing computations once a finite symbolic message to be processed becomes rigidly connected (as it is actually the fact, Section X C) to its infinitely long context, *non-symbolically* presented as a real-world physical recall/recognition device or brain circuitry. In other words, any super-Turing computer processing a meaningful symbolic message might indeed be considered as a special-purpose regular Turing computer but running in the unique, specific to it *real-world environment that is also a part of computational process and computational device*. A Turing-type computer is inevitably a part of a super-Turing computer that is a *qualitatively distinct* computational machine because it combines symbolism and physical reality to process particular infinite-in-length symbolic strings or real-valued or continuous numbers for a finite time period. Another essential innovation is the use of *mirror* super-Turing machines, in order that communicators would be able to understand (correctly decode) a finite *meaningful* symbolic message addressed from one of them to another (see also Section X B3). BSDT PL super-Turing computations are not purely symbolic "tautological" transformations and super-Turing computers are not a set of connected elementary discrete-logic devices for doing these transformations – *both of them* are "an inseparable mix" of symbolism and reality. It is worth noting, that is also the reason why referring to this combined symbolism-reality computational method, we prefer the terms "primary language" and "language of primary thoughts" over "semantic mathematics" and "mathematics of meaningful computations."

E. Memory Performance without Knowing Memory Records, Continuity-discreteness Unity and Uncertainty

Since the BSDT PL employs for naming the things to be named a non-Gödelian arithmetization by natural numbers x_k^i (Section VII A) and since these natural numbers are *randomly* chosen from their finite range of values (Section V D4), particular random choice of a name does not matter and the following effect was predicted [17]. By empirical examination of an ASM hierarchy/neural subspace [28], [73] that generates the meaning of a trace x_k^i (Figure 8), all the parameters describing the ASM(x_k^i) may successfully be found but the content of the x_k^i – specific given the *i* randomly-established arrangement of its ±1 components – will always remain unknown. If it is, then, for example, the content of a particular given-length memory record does not affect memory performance and *can not empirically be found*. This rather surprising prediction [17] has well been corroborated by numerical BSDT PL analysis [66] of empirical receiver operating characteristics, ROCs (functions providing memory performance).

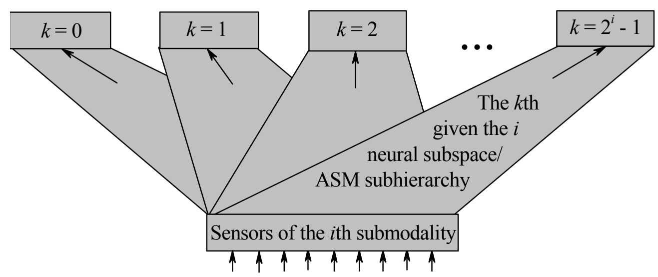

Figure 8. Neural subspaces/ASM subhierarchies generated the meanings of 2^i of given the *i* words x_k^i. All the subhierarchies (trapeziums) of the *i*th submodality are fed by impulses generated by the same set of sensors (lower rectangle) and the *k*th subhierarchy produces the *k*th pattern (arrow) of impulses to the *k*th apex ASM (ASM(x_k^i), the *k*th upper rectangle) learned to store and recognize the x_k^i. The correspondence between the *k*th given the *i* name (the arrangement of components of the x_k^i stored in the *k*th upper rectangle) and the thing it names is fixed but *randomly* established and empirically can not be found [66].

In [66] discrete-valued memory-for-meaningful-words ROCs measured in healthy humans and patients with brain disorders [80] were fitted by the BSDT. For this purpose, words and networks storing these words were presented as binary vectors x_k^i and respective BSDT ASMs. These ASMs are devoted to recall/remember/recognize the only memory trace specific to it [30]; for example, the ASM(x_k^i) serves the x_k^i. In Figure 8, the ASM(x_k^i) is an apex ASM of the *k*th ASM subhierarchy/neuron subspace generating the inputs to this ASM and giving the x_k^i its meaning (this scheme is called a semi-representational memory model [28], [73]). The size *N* of the network storing the x_k^i (in Figure 8, $N = i$), the intensity *q* of the cue used in the process of a memory trace retrieval, the preferred rate j of the expected in experiment decision confidence, and the arrangements of components of vectors x_k^i (in Figure 8, they are enumerated by the index k) were used as fitting parameters. As a result of fitting, empirical discrete-valued ROCs for healthy subjects and patients with brain disorders were numerically reproduced by BSDT calculations (Figure 3 in [66]) and values of parameters N, q, and j were successfully found (Table 1 in [66]) *without* any reference to arrangements of components of vectors x_k^i. Thus, in full accordance with the BSDT PL prediction, complete BSDT description of performance of the memory-for-meaningful-words can indeed be achieved without knowing the memory records. Consequently, the ideas of non-Gödelian BSDT PL arithmetization by natural numbers (Section VII A) and, simultaneously, BSDT semi-representational memory model [28], [73] have indeed numerically and empirically been substantiated.

In spite of the BSDT's essential discreteness, it contains one *continuous* physical/physiological parameter, namely the neuron triggering threshold θ [66], [81]. In the course of ROC fitting [66], all the values of the θ from their jth finite-in-width range $\Delta\theta_j$, $\theta \in \Delta\theta_j$, are transformed into the only *integer* value of the decision confidence j from its range $0 \leq j \leq N + 2$. Reverse transformation of this value of j into a certain $\theta \in \Delta\theta_j$ that just generated the j is impossible and this fact is the BSDT implementation [66] of the predicted by the BSDT PL continuity-discreteness unity and uncertainty (Section VI B). This prediction and its BSDT implementation are well supported by the empirical discovery of irremovable spike onset potential (neuron triggering threshold) variability up to 10 mV [82]. In terms of neuroscience this variability is explained by fluctuating synaptic currents and by inherent statistics of the opening cell channels while in BSDT terms it is the width of the $\Delta\theta_j$ in voltage units.

XI. Examples and Perspective BSDT PL Applications

Practical examples of BSDT PL meaningful computations could help to better understand their features and the perspectives of their further applications.

A. BSDT PL Computations, the Concept and the Manual

The BSDT PL's *concept* of semantic computations is surprisingly simple because it recommends, before the beginning of calculations, *to know* complete formal description of reality. At the same time, it is surprisingly complex because it recommends, before the beginning of calculations, *to find* complete formal description of the reality. As such a description (the context) is by definition of an infinite length, these recommendations can of course never be fulfilled completely. That is, the main problem of meaningful computations is the incompleteness of knowledge of their context or, in other words, the incompleteness of available formal or "mathematical" descriptions of reality. As soon as such a description has been found and fixed, BSDT PL semantic computations are reduced to usual Turing computations and could easily be performed, e.g., [28], [65], [66].

In addition to these general recommendations, the BSDT PL gives also *a manual* for meaningful computations. It is based on the BSDT (a theory providing the best encoding-decoding rules [31, 73] for binary finite-dimensional vectors x_j^i damaged by replacing binary noise [27]) and the technique of BSDT ASMs (abstract selectional machines [30] implementing the BSDT encoding-decoding rules or BSDT PL inference rules). Given the context, the BSDT implements the main distinct features of meaningful BSDT PL computations: 1) the discreteness of all the computations with finite binary vectors x_j^i [65], [81], 2) the uniqueness of the vector x_j^i a particular ASM is devoted to process in the best way [30], and 3) the ability of each ASM to generalize even from a single example [73]. The first of these features leads to a fundamental discreteness of all BSDT PL computational predictions found at precisely fixed context. The second feature generates one-memory-trace-per-one-network network learning paradigm. The third feature ensures the BSDT PL's tolerance to damages and noise and its capability of coping with "effective stochasticity" of an agent's permanently changing environment.

For the study of meaningful information exchanges, their actual context should empirically be estimated with maximal possible accuracy. This is not a trivial problem and it is a subject of intensive research, e.g., [83], [84], [85]. Results available in this field are so far insufficiently rich because the required measurement methods remain till now in the state of development.

B. Meanings of Traditional Solutions of Mathematical Problems of Science and Practice

Besides the axioms, any formal axiomatic system, FAS, comprises symbolic descriptions of all its theorems and inference rules. In the BSDT PL, an FAS (e.g., ZFC) is represented as an *infinite fraction* of meaningless finite binary strings, x_j^i, that are the affixes of meaningful strings $c_{xi}x_j^i \in S_{cx0}$ (Section V B). For this reason, any FAS computations are also BSDT PL computations and numerous already available computational results, e.g., in physics or biology may be treated as examples of BSDT PL computations performed given a context defined *formally and informally*.

A separate infinite BSDT PL string $c_{xi}x_j^i$ that gives a meaning $M(x_j^i) = c_{xi}x_j^i$ to a finite symbolic message x_j^i (it may be, e.g., a physical formula written in binary notations) includes an infinite description c_{xi} of the FAS needed to derive this formula and of the physical problem that gives this formula a physical sense. For professionals (P) and laypersons (L), this formula has different meanings we denote as $M_P(x_j^i)$ and $M_L(x_j^i)$, respectively: $M_P(x_j^i) = c_{xi}(P)x_j^i = c_{xi}(IP)c_{xi}(FP)x_j^i$ and $M_L(x_j^i) = c_{xi}(L)x_j^i = c_{xi}(IL)c_{xi}(FL)x_j^i$ where finite-in-length strings $c_{xi}(FP)$ and $c_{xi}(FL)$ represent the formal knowledge (F) and infinite-in-length strings $c_{xi}(IP)$ and $c_{xi}(IL)$ represent informal knowledge (I) about the formula and the problem of interest. Formal knowledge can be found in books or any other relevant texts, informal knowledge can only be acquired from individual experiences of professionals and laypersons, respectively.

Since professionals and laypersons have essentially different backgrounds in a particular knowledge domain, $c_{xi}(IP)c_{xi}(FP) \neq c_{xi}(IL)c_{xi}(FL)$, they understand the meaning of the x_j^i in a different way. Since $c_{xi}(FP)$ and $c_{xi}(FL)$ are finite and may explicitly be specified, they may be compared explicitly (e.g., by the grading of school exams). Since $c_{xi}(IP)$ and $c_{xi}(IL)$ are one-way infinite and essentially unspecified, it is impossible to compare them explicitly. We may only know that they are of different meaning complexities (Section V C) and have somewhere "in the past" a common infinite initial part. If to remember our phenomenology formalization (Figure 2) then it becomes clear that informal or "implicit" knowledge is presented in the brain as non-symbolical properties of real-brain devices serving this knowledge. Informal character of implicit knowledge also indicates the crucial role of the teacher and educational environment (the supervisor and research environment) for acquiring knowledge. Consequently, as an important source of informal knowledge, the teacher/ supervisor can never be excluded from the process of teaching/research training.

Formal or "explicit" knowledge – strings $c_{xi}(FP)$ and $c_{xi}(FL)$ – is not the content of books we have read in a school or university but their individual internal symbolic representation that, in terms of the BSDT PL, may be different in different minds. Informal or implicit knowledge – strings $c_{xi}(IP)$ and $c_{xi}(IL)$ – we acquire from our *personal* experiences under different teachers/supervisors in different educational/research environments and, consequently, it is also different for each of us. For these reasons, different professionals and different laypersons read the same books but understand them differently. In particular, for different professionals which we call P_1 and P_2, the formula x_j^i always has to an extent different meanings, $M_{P1}(x_j^i) \neq M_{P2}(x_j^i)$, and, consequently, even in so-called "exact" sciences a vagueness of meanings of their formal results is unavoidable and can not completely be excluded. In other words, for traditional FAS computations their context and, consequently, their meanings for different peoples can never *precisely* be fixed. Hence, in this case, the BSDT PL may only approximately be applied: its inherent discreteness is masked by the vagueness of knowledge each of us have about the context of ZFC computations. At the same time, as all humans (professionals and laypersons) are of the same species, their knowledge is internally represented by infinite strings of the same meaning complexity and, consequently, all of us can understand anything that understands anyone else on the condition of course that beforehand we were equally prepared/trained (see also Section X D).

Relationships, which were just described, between formal/informal knowledge and traditional computations, draw our attention to the fact that any manipulations with numbers will be meaningless until their giving-the-meaning context is added and fixed.

C. Processing Meaningful Memory Records and Meaningful Images Given their Precisely Fixed Contex

A situation may of course be conceived when the context of different symbolic messages is completely, bit-by-bit the same without any reservations. *For particular animal/human*, it may be, e.g., the case of members, $c_{xi}x_j^i$, of the same category of names, $C(x_j^i)$. If the context of names x_j^i, c_{xi}, is *precisely fixed* then the inherent discreteness of the BSDT PL should be visible as inherent discreteness of respective empirical data and these inherently discrete data should successfully be described by the inherently discrete BSDT PL computations. The main obstacle is the need of discovering such inherently discrete natural phenomena and developing a methodology of research that will not hide their discreteness.

As has been demonstrated in [28], [65], [66] all the mentioned conditions can be satisfied. As a result, we have already three particular *examples of complete successful application of given the context discrete-valued BSDT PL computations* to account for practically important cognitive (where the role of mind is essential) phenomena in humans. They are in particular 1) judgment errors in cluttered environments [65], 2) remembering/retrieving the words from a memory for meaningful words [66], and 3) recognition of meaningful images (human faces) by healthy humans [28]. In the first of these cases, with the help of the BSDT, the data measured in rating experiments when healthy subjects identify target stimuli in a cluttered visual environment, confound them with competing stimuli, and demonstrate high confidence of their erroneous decisions were quantitatively explained (Figure 4 in [65]); in the second case, memory-for-meaningful-words ROCs measured in healthy humans and patients with brain disorders were quantitatively described by the BSDT and memory-for-meaningful-words parameters were found (Figure 3 and Table 1 in [66]); in the third case, psychometric functions measured in human face recognition experiments were reproduced by the BSDT keeping the Neyman-Pearson objective (Figure 5 in [28]).

In each of these examples, BSDT discrete-valued numerical analysis has been applied to fitting empirical data measured by traditional techniques and analyzed by the authors of original publications [80], [86], [87] using traditional continuous computations. The authors of these publications did not recognize the discreteness of their results, in particular, because of the essential *continuity* of mathematical models they employed for empirical data analysis. We take the opposite view of these models motivated by the BSDT PL and its hypothesis of concurrent infinity. Namely, results of cognitive experiments found at the precisely defined context are to be *discrete* because in such a case the continuous/real-valued component of meaningful messages (infinite-in-length context or the set of real-brain circuits/devices involved in serving the cognitive tasks) is strictly the same, fixed, excluded from explicit consideration and "invisible" in practice as a result. Indeed, if a given person recalls different meaningful words or recognizes different meaningful images then he/she is dealing with different finite symbolic messages x_j^i *given* bit-by-bit the same context c_{xi} defined by this person's unique previous experience that shapes his/her unique mind, relevant to a particular problem. This mind or particular set of real-brain learned circuits or particular neuron subspace or particular context c_{xi} is surely the same in all the tasks of the same type this mind (person) currently performs (cf. Figure 8). On the other hand, it is the discreteness of empirical data that is a manifestation of certainly definite meanings of symbolic messages, involved in context specific cognitive tasks. In other words, meaningfulness of messages to be processed and their precisely known context are simultaneously the source of this type of data. The continuity of observed cognitive performance indicates the vagueness of the context's estimation (inability of keeping the fixed context) that may be caused by irrelevant choice or inaccurate use of measurement protocols. It is supposed, if meanings are fixed and exactly communicated then the context is completely the same and communication (encoding/ decoding) performance should surely be discrete. As the fitting of empirical data measured in these three different types of cognitive experiments demonstrates [28], [65], [66], some popular study protocols seems to produce discrete-valued data in cognitive sciences though, to be convinced, new BSDT fitting results and some control experiments are surely required [66].

The need to keep the same context to ensure accurate meaningful communication/computation is rather well known, e.g., [83], [84], [85]. In cases where this demand is accurately satisfied, methods developed by other authors coincide with the BSDT PL sometimes almost literally. For example, what is called in [88] "meaning-generating capacity" of a complex dynamic system, namely "the proportion between the size m of the set of final attractor states and the size n of the set of all initial states of a system, i.e., $MC = m/n$" is in BSDT terms the ASM(x_k^i) probability of correct decoding given the size of the network $N = i$, intensity of cue q, and decision confidence rate j. This probability is denoted as $P(N, q, j)$ or $P(N, q, F_j)$ (F_j is false alarm probability given the j) and was already successfully used to analyze the results of real cognitive experiments [28], [65], [66]. The main distinction between the MC and the $P(N, q, j)$ is that the former is defined given the finitely estimated context whereas the latter is fixed given an infinitely defined context. The fact that the $P(N, q, j)$ is perfect [31], [89] and may even analytically be found [89] is secondary with respect to the context infinity.

D. *A Lecturer and Students in a Lecture Room*

Let us consider a lecturer who intends to deliver students the meaning of a physical formula as he/she understands it. At the beginning of a lecture, he/she and his/her students have different background knowledge of the formula of interest and students can not correctly understand its meaning. The lecturer's aim is to give them a piece of additional knowledge and, in this way, to equalize, for all of them, the context needed to equally understand the formula's meaning. At the end of the lecture, for the lecturer and for his/her students, infinite BSDT PL strings describing this specific knowledge should become bit-by-bit equivalent not only "in the past" but also "in the present", and the formula's meaning should be understood by all the parties in the same way. If it is not for any reason, a misunderstanding arises.

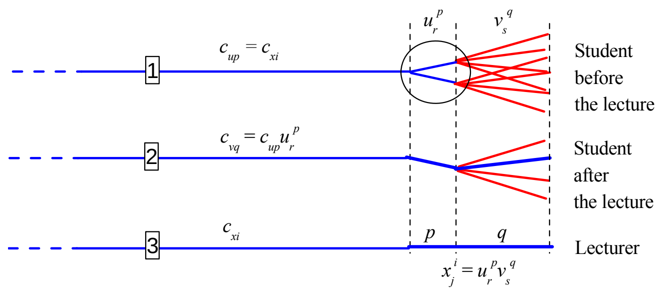

Figure 9. BSDT PL model of information exchange between the lecturer (line 3) and students before (line 1) and after (line 2) the lecture. The amount of information of interest is q bits, the "gap" (it is circled) between the knowledge of the lecturer and the knowledge of students equals p bits (examples for $q = 2$, $p = 1$; $i = q + p$). The fan of line segments attached to line 1 represents given the context c_{up} the set of possible meanings of the formula of interest in students before the lecture. The fan attached to line 2 gives the set of possible meanings after the lecture (after bridging the knowledge gap), thick line segments coincide bit-by-bit with the BSDT PL representation of the lecturer's knowledge (the thick fraction of the line 3). Other designations as in Figure 6.

In Figure 9 some aspects of the process of teaching and learning are presented in BSDT PL terms. It means, we ignore so far the fact that the lecturer and students use a natural language for their communication and we are not interested in mechanisms of translation of a natural language into the BSDT primary language and vice versa. As the lecturer and students are of the same species, they all use the same primary language (we suppose, BSDT PL) and their meaningful words are of the same meaning complexity. Such a representation is person-dependent and natural-language-independent. Line 1 in Figure 9 represents the *i*th submodality (a particular set of brain circuits that are ready to be changed or a plasticity area, cf. Figure 8) allocated in the brain of a student before the lecture to write, store and then retrieve particular information this student intends to acquire from the lecture. Line 2 gives the same for a successful student after the lecture; line 3 represents the lecturer's same brain area. One-way infinite fractions of these lines, $c_{xi} = c_{up}$ (from the left to the first vertical dashed line), designate a description of everything that is in common in the lecturer and his/her students, from genetic code to textbooks they have read. Composite vector $x_j^i = u_r^p v_s^q$ describes new knowledge the lecturer intends to deliver. This vector can be divided into two constituents one of which (v_s^q) describes the physical formula of interest and the other (u_r^p) describes the additional information ("local context") needed to connect the v_s^q to already available background knowledge, $c_{xi} = c_{up}$. A vector u_r^p or, more accurately, the form u^p (Section V B) may be treated as p-bits-in-width "gap" between the knowledge of the lecturer and the knowledge of students that should be bridged before the students would be able to understand the formula. As a result of teaching, a student's area of plasticity responsible for acquiring this specific knowledge changes and reduces 2^i of different ways of understanding the message the lecturer communicates to the only one, the same as of the lecturer. It means, all the parties have now the same BSDT PL internal representation of the knowledge of current interest (thick line segments in lines 2 and 3) and equally understand it.

E. Non-syntactic and Non-language Communication by Basic Bihaviors

In the previous example (Section XI D), non-syntactic messages represent a very small fraction of the general flow of information. Among them it may be, e.g., the facts that the lecturer is walking when he/she gives his/her talk. This non-syntactic and even non-language message (bodily signal) is effortlessly understood by everyone who is in the room because all people are members of the same species, have the same *innate* bodily infrastructure and the same basic behaviors (Section X B1) developed from their same *innate* behavioral reflexes, e.g., [69], [70], [71] in the course of human learning and natural ageing (maturation) in a mostly common environment. As a result, adults or infants of relevant ages have *common mirror neuron systems* (Section X B3 and D) to produce and perceive/understand their basic behaviors included, for example, walking. For this reason, humans/animals produce and perceive such non-syntactic non-language messages originated from their basic behaviors automatically, with practically no chance of misunderstanding. For all given the species animals (or humans), meanings of their basic behaviors are given by practically equivalent brain circuits or infinite BSDT PL strings of the same length (to remind, possible the strings' distinctions are inessential because of BSDT ASM tolerance to damages and noise [30], [73], Section X B3).

The role of mirror neuron systems for understanding the actions and possibly the intentions of others, e.g., [90], [91], [92] and their role in evolutionary language development are rather well recognized [93], [94] and even to a degree studied by the method of computational modeling, e.g., [95], [96]. In these publications, the importance of training the innate brain structures for a design, on their ground, of mirror neuron systems and the importance of mirror systems for mimicking actions and language production are in particular emphasized. At the same time, contrary to the BSDT PL assumption, Michael Arbib and his colleagues suppose [93] - [96] that super-Turing computability is not relevant to brain computations. Such an attitude seems indeed rather natural while we are dealing, as it is usually the case, with *meaningless* computations only and do not pay attention to their meanings. But as soon as meanings become essential super-Turing computability escapes from the shadow of conventional Turing computations and becomes crucially important. It is what is the case for the BSDT PL because it does imply that super-Turing computability, as an indispensable part of animal/human communication process, maintains mechanisms of doing all the meaningful actions the brain serves (Section X B3 and D), including all kinds of non-language and language meaningful information exchange.

F. Natural Languages and Consciousness, Intuition, Free Will and Creativity

One of distinctive features of the BSDT PL is that truth values of its names are always in the norm true (Section VIII). That is why it serves so well to as a primary language for maintaining an animal's ongoing internal activity. For the same reason, it can serve as a "source language" whose meaningful words (an animal's psychological states) may next be translated into vocal, gesture, etc tokens of a more elaborate symbolic communication system needed to support information exchange between animals of a group. The more complicated the group's sociality, the more complicated communication system is required to support it, and vice versa. Since among other animals humans do have most complicated sociality, human natural languages are to be most complicate and elaborate. The BSDT PL may be used as a basis for the construction of such "secondary" [26] languages whose capacities may be up to the level of human natural language capacity. If so, semantics and syntax of natural languages should be based on semantics and syntax of the BSDT PL and should be implemented by mechanisms of (and innate brain structures for) the translation of words/sentences of the primary language into words/ sentences of a secondary language. In that sense the BSDT PL is a counterpart or a precursor to what is known as Noam Chomsky's "universal grammars," e.g., [97], [98].

The BSDT PL phenomenology formalization literary equates one-way infinite binary strings and animal psychological states or subjective experiences/qualia. It also represents given the context computations with finite binary meaningful strings as operations with animal subjective experiences. In other words, the BSDT PL solves the "hard" problem of consciousness [33] (the quest for a description of subjective experiences) as a whole and at once, simply by *postulating* the logically strict BSDT PL definition of qualia (Sections III and V B). For this reason, the BSDT PL is actually a theory of subjectivity (meaning, feeling, perception) or *a theory of the subconscious*. The problem remains to apply this theory to solving particular practical consciousness problems, as it has, e.g., been done in the case of BSDT atom of consciousness model, BSDT AOCM [32]. In particular, the mentioned above problem of translating the primary language into a secondary one may also be treated as the problem of translating the subconscious *served by super-Turing computations* into the conscious *served by Turing computations*. In both cases, the process of translation should inevitably be based on so far unspecified mechanisms of *intuition, free will* and *creativity*.

XII. CONCLUSIONS

The BSDT PL is based on the hypothesis of concurrent infinity and its phenomenology formalization (Sections I to IV). It provides what is called a "paradigm shift" [36]: a possibility to equate the items of such usually incommensurable domains as the symbolism and reality, to define strictly indefinable in traditional mathematics notions of meaning and subjectivity, and to perform explicitly given the context meaningful computations. Such computations are an inherent mix of symbolism and specific to it reality implemented by a qualitatively novel computational device – a super-Turing computer with infinite inputs implemented in an animal's brain as a system of mirror neurons (Section X D). The range of perspective BSDT PL applications covers everything where meanings are important or, in other words, everything we, humans, may be interested in. In this sense it may be "a theory of everything". As soon as meanings become inessential, the BSDT PL is reduced to traditional ZFC mathematics. Available empirical and computational results support this view (Sections X and XI).

BSDT PL provides a framework that is sufficient to perform principal semantic computations and based on them communication without syntax. BSDT PL seems also to be sufficient to explain the computational part of intelligence of animals of poor sociality and, consequently, to design the computational part of intelligence of artificial devices (e.g., robots) or computer codes mimicking the behavior of such animals. At the same time, the BSDT PL is unable to symbolically explain the mechanism of splitting its composite words (sentences) into focal and fringe constituents (Section VI B2) and, consequently, of directing an animal's attention to a particular thing – we hope it may be done by methods beyond the discrete BSDT formalism. To explain/reproduce the "attentive" part of animal intelligence in a biologically-plausible way and to design the "attentive" part of the intelligence of intelligent robots, analog (e.g., wave-like) computational methods similar to those that are used in real brains are most probably required.

Contrary to traditional formal languages, e.g., [98], [99] that are in end the products of traditional ZFC mathematics, the BSDT PL is a consistent and complete (Sections VI to VIII) calculus of finite binary strings (spike patterns or "symbols") with *infinitely* defined contexts. It is based on 1) the new infinity hypothesis and its phenomenology formalization (Sections I to IV) providing the technique of super-Turing (semantic) computations with infinite binary strings that share their infinite initial part and 2) the BSDT [25] and its ASMs [30] providing a technique for the best encoding/decoding in binary finite-dimensional spaces [27] and implementing BSDT PL inference rules. BSDT PL is the simplest language of its kind and has great potential for designing the adequate models of higher-level languages, including in perspective the natural languages of humans. At the same time, meaning ambiguity of BSDT PL names of different meaning complexity that has been established as their fundamental property (Section IX) raises many intriguing problems to be solved in the future.

The BSDT PL describes a way for the communication of meanings of symbolic messages by means of basic animal behaviors (Sections X B1 and XI E) that could represent the behavioristic part [72] of more complex adaptive animal behaviors. For animals of the same and, in many cases, of relative species, thanks to their mirror neurons, e.g., [77] - [79] and common "bodily infrastructure" [34], it is intelligible without any efforts. For animals with most primitive sociality (including human infants) or for their *artificial* counterparts, a version of the discrete BSDT PL formalism may serve as an *exhaustive* but *incomplete* (Section VI B3) set of tools needed for their routine communication. How the primary language generates secondary (natural) languages and consciousness [32] is the problem of future research.

Following Rudolf Carnap [100, p. 204] let us finish this article by a quotation from Bertrand Russell (his term "denotation" may here be understood as "meaning"): "Of many other consequences of the view I have been advocated, I will say nothing. I will only beg the reader not to make up his mind against the view—as he might be tempted to do, on account of its apparently excessive complication—until he has attempted to construct a theory of his own on the subject of denotation. This attempt, I believe, will convince him that, whatever the true theory may be, it cannot have such a simplicity as one might have expected beforehand" [101, p. 518].

ACKNOWLEDGMENT

I thank the participants of the conference INTELLI-2012 for their comments on my oral presentation and Prof. Michael Arbib for his comments to its published version. I am grateful to my family and my friends, especially to my son Dr. Mykhaylo Gopych, for their help and support.

REFERENCES

[1] P. Gopych, "Primary language for semantic computations and communication without syntax," in INTELLI 2012 – The First International Conference on Intelligent Systems and Applications, P. Lorenz and P. Dini, Eds. Chamonix, France, 2012, pp. 47-53.

[2] G. H. Hardy. A mathematican's apology. Cambridge, England: Cambridge University Press, 1941.

[3] L. A. White, "The locus of mathematical reality: An apological footnote," in The World of Mathematics, vol. 4, J.R. Newman, Ed. New York: Simon and Shuster, 1956, pp. 2348-2364.

[4] A.A. Fraenkel, Y. Bar-Hillel, and A. Levy. Foundations of set theory, 2nd ed. Amsterdam: Elsevier, 1973.

[5] W. Sieg, "Hilbert's program: 1917-1922," Bull. Symb. Logic, vol. 5, March 1999, pp. 1-44, http://www.jstor.org/stable/421139.

[6] K. Gödel. On formally undecidable propositions of Principia Mathematica and related systems, English transl. by B. Meltzer. New York: Dover, 1992. Originally published Kurt Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I," Monatshefte für Mathematik and Physik, vol. 38, 1931, pp. 173-198.

[7] A. Turing, "On computable numbers, with an application to the Entscheidungsproblem," Proc. London Math. Soc. (2) vol. 42, 1936, pp. 230-265; vol. 43, 1937, pp. 544-546; in Collected works of Alan Turing. Mathematical logic, R.O. Gandy and C.E.M. Yates, Eds. Amsterdam: Elsevier, 2001, pp, 18-54.

[8] H. Poincaré. Science and method. London: Thomas Nelson and Sons, 1914.

[9] S.C. Kleen. Introduction to metamathematics. Amsterdam: North-Holland Publ., 1952.

[10] A. Heyting. Intuitionism. An introduction. Amsterdam: Noth-Holland Publ., 1956.

[11] J.R. Lucas, "Minds, machines and Gödel," Philosophy, vol. 36, Apr.-Jun. 1961, pp. 112-127.

[12] R. Penrose. The emperor's new mind. New York - London: Pinguin Books, 1991.

[13] R. Penrose. Shadows of the mind: a search for the missing science of consciousness. Oxford: Oxford University Press, 1994.

[14] J.R. Searle. The mistery of consciousness. New York: New York Review Book, 1997.

[15] W.T. Fitch. The evolution of language. Cambridge, UK: Cambridge University Press, 2010.

[16] H. Weyl, Philosophy of mathematics and natural scienves, 4th ed. Prinecton: Princeton University Press, 1959.

[17] P. Gopych, "On semantics and syntax of the BSDT primary language," in KDS 2010, Information models of knowledge, K. Markov, V. Velychko, and O. Voloshin, Eds. Kiev-Sofia: ITHEA, 2010, pp. 135-145, http://foibg.com/ibs_isc/ibs-19/ibs-19-p15.pdf <retrieved: November, 2012>.

[18] E.O. Wilson. Conscilience: the unity of knowledge. New York: Vintage Books, 1999.

[19] H. Maturana and F. Varela. Autopoiesis and cognition. Dordrecht, Holland: D. Reidel Publ., 1980.

[20] L. Margulis. Symbiosis in cell evolution, 2nd ed. San Francisco: Freeman, 1993.

[21] C.B. Pert, M.R. Ruff, R.J. Weber, and M. Herkenham, "Neuropeptides and their receptors: a psychosomatic network," The Journal of Immunology, vol. 135 (Suppl.), Aug. 1985, pp. 820-826.

[22] M. Minky. The society of mind. New-York: Simon and Schuster, 1988.

[23] P.O.A. Haikonen. Robot brains: circuits and systems for conscious machines. Chichester, UK: John Wiley and Sons, 2007.

[24] D. Dennett. Consciousness Explained. London: Penguin Books, 1993.

[25] P.M. Gopych, "Elements of the binary signal detection theory, BSDT," in New research in neural networks, M. Yoshida and H. Sato, Eds. New York: Nova Science, 2008, pp. 55-63.

[26] J. von Neumann. The computer and the brain. New Haven: Yale University Press, 1956.

[27] P. Gopych, "BSDT multi-valued coding in discrete spaces," in CISIS-08, ASC, vol. 53, E. Corchado, R, Zunino, P. Gastaldo, and Á. Herrero, Eds. Berlin-Heidelberg: Springer, 2009, pp. 258-265, doi: 10.1007/978-3-540-88181-0_33.

[28] P. Gopych, "Biologically plausible BSDT recognition of complex images: the case of human faces," Int. J. Neural Systems, vol. 18, Dec. 2008, pp. 527-545, doi: 10.1142/S0129065708001762.

[29] G.M. Edelman, "Naturalizing consciousness: a theoretical framework," Proc. Natl. Acad. Sci., vol. 100, Apr. 2003, pp. 5520–5524, doi: 10.1073/pnas.0931349100.

[30] P. Gopych, "Minimal BSDT abstract selectional machines and their selectional and computational performance," in IDEAL-07, LNCS, vol. 4881, H. Yin, P. Tino, E. Corchado, W. Byrne, and X. Yao, Eds. Berlin-Heidelberg: Springer, 2007, pp. 198-208, doi:10.1007/978-3-540-77226-2_21.

[31] P.M. Gopych, "Foundations of the neural network assembly memory model," Int. J. Comput. Res., vol. 13, 2004, pp. 103-166; in Leading-edge computer sciences, S. Shannon, Ed. New York: Nova Science, 2006, pp. 21–84.

[32] P. Gopych, "BSDT atom of consciousness model: the unity and modularity of consciousness," in ICANN-09, LNCS, vol. 5769, C. Alippi, M.M. Polycarpou, C. Panayiotou, and G. Ellinas, Eds. Berlin-Heidelberg: Springer, 2009, pp. 54-64, doi: 10.1007/978-3-642-04277-5_6.

[33] D. Chalmers. The conscious mind. Oxford: Oxford University Press, 1996.

[34] N.J. Enfield, "Without social context?" Science, vol. 329, Sep. 2010, pp. 1600-1601, doi: 10.1126/science.1194229.

[35] E. Husserl. The idea of phenomenology, Enlish transl. by L. Hardy. Dordrecht: Kluwer Academic Publ., 1907/1999.

[36] T.S. Kuhn. The structure of scientific revolutions, 3rd ed. Chicago-London: Universithy of Chicago Press, 1996.

[37] L. Wittgenstein. Tractatus logico-philosophicus. London - New York: Routledge, 1922/1974.

[38] T. Ord, "The many forms of hypercomputations," App. Math. Comp., vol. 178, Jul. 2006, pp. 143-153, doi: 10.1016/j.amc.2005.09.076.

[39] W.V. Quine. Set theory and its logic. Cambridge, MA: Harvard University Press, 1969.

[40] L. Pozsgay, "Liberal intuitivism as a basis of set theory," Proc. Sym. Pure Math., vol. 13, part I. Providence, Rhode Island: AMS, 1971, pp. 321-330.

[41] D. Perrin and J.-É. Pin. Infinite words. Amsterdam: Academic Press, 2004.

[42] K. Mainzer. Thinking in complexity, 5th ed. Berlin: Springer, 2007.

[43] V. McGee. Truth, vagueness, and paradox. Indianapolis: Hackett, 1991.

[44] S. Ullman, "Object recognition and segmentation by a fragment-based hierarchy," Trends Cogn. Sci., vol. 11, Feb. 2007, pp. 58-64, doi: 10.1016/j.tics.2006.11.009.

[45] C. Shannon, "A mathematical theory of communication," Bell Syst. Techn. J., vol. 27, 1948, pp. 379-423, 623-656.

[46] M. Li and P. Vitanyi. An introduction to Kolmogorov complexity and its applications, 2nd ed. Berlin-Heidelberg: Springer, 1997.

[47] M. Gell-Mann and S. Lloyd, "Information measures, effective complexity, and total information," Complexity, vol. 2, Sep.-Oct. 1996, pp. 44-52, doi: 10.1002/(SICI)1099-0526.

[48] S. Kripke. Naming and necessity, rvd. ed. Oxford: Basil Blackwell, 1990.

[49] K. Gödel. The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory. Princeton, NJ: Princeton University Press, 1940.

[50] P.J. Cohen, "The independence of the continuum hypothesis I," Proc. Natl. Acad. Sci. USA, vol. 50, Dec. 1963, pp. 1143-1148.

[51] P.J. Cohen, "The independence of the continuum hypothesis II," Proc. Natl. Acad. Sci. USA, vol. 51, Jan. 1964, pp. 105-110.

[52] D. Hilbert, "Mathematical problems. Lecture delivered before the International congress of mathematicians at Paris in 1900," English transl. by M. W. Newson, Bull. Amer. Math. Soc., vol. 8, 1902, pp. 437–479.

[53] K. Whittingstall and N.K. Logothetis, "Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex," Neuron, vol. 64, Oct. 2009, pp. 281-289, doi: 10.1016/j.neuron.2009.08.016.

[54] B.S.W. Ng, N.K. Logothetis, and C. Kayser, "EEG pase patterns reflect the selectivity of neuron firing," Cerebral Cortex, Feb. 2012, doi: 10.1093/cercor/bhs031.

[55] D.R. Bach and R.J. Dolan, "Knowing how much you don't know: a neural organization of uncertainty estimates," Nat. Rev. Neurosci., vol. 13, Aug. 2012, pp. 572-586, doi: 10.1028/nrn3289.

[56] G. Thut, A. Nietzel, S.A. Brandt, and A. Pascual-Leone, "Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection," J. Neurosci., vol. 26, Sep. 2006, pp. 9494-9502, doi: 10.1523/jneurosci.0875-06.2006.

[57] T.P. Zanto, M.T. Rubens, A. Thangavel, and A. Gazzely, "Causal role of prefrontal cortex in top-dowm modulation of

visual processing in working memory," Nat. Neurosci., vol. 14, May 2011, pp. 656-661, doi: 10.1038/nn.2773.

[58] A. Zénon and R.J. Krauzlis, "Attention deficits without cortical neuronal circuits," Nature, vol. 489, Sep. 2012, pp. 434-437, doi: 10.1038/nature11497.

[59] A. Smolyanskaya and R.T. Born, "Attention is more than meets the eye," Nature, vol. 489, Sep. 2012, pp. 371-372, doi: 10.1038/489371a.

[60] K. Gödel, "Remarks before Princeton bicentennial conference on problems of mathematics," in The undecidable, M. Davis, Ed. New York: Raven Press, 1965, pp. 84-88.

[61] C.S. Calude and G.J. Chaitin, "Randomness everywhere," Nature, vol. 400, Jul. 1999, pp. 319-320, doi: 10.1038/22435.

[62] G. Chaitin. The limits of mathematics. Singapore: Springer, 1998.

[63] G. Lolli, "Peano and the foundations of arithmetic," in Giuseppe Peano between mathematics and logic, F. Skot, Ed. Milan: Springer, 2011, pp. 47-67.

[64] W.V. Quine. Pursuit of truth. Cambridge, MA: Harvard University Press, 1992.

[65] P.M. Gopych, "Performance of BSDT decoding algorithms based on locally damaged neural networks," In IDEAL-06, LNCS, vol. 4224, E. Corchado, H. Yin, V. Botti, C. Fyfe, Eds. Berlin-Heidelberg: Springer, 2006, pp. 199-206, doi: 10.1007/11875581_24.

[66] P. Gopych and I. Gopych, "BSDT ROC and cognitive learning hypothesis," in CISIS-10, AISC, vol. 85, Á. Herrero, E. Corchado, C. Redondo, Á. Alonso, Eds. Berlin-Heidelberg: Springer, 2010, pp. 13-23, doi:10.1007/978-3-642-16626-6_2.

[67] A. Tarski. Logic, semantics, metamathematics, 2nd ed. Oxford: Oxford University Press, 1935/1983.

[68] B. Russell. Principles of mathematics. London - New York: Routledge, 1903/2010.

[69] J.L. Pressler and J.T. Hepworth, "Newborn neurologic screening using NBAS reflexes," Neonatal network, vol. 16, Sep. 1997, pp. 33-46.

[70] J. Scott and M. Rosser, "The grasp and other primitive reflexes," J. Neurol. Neurosurg. Psychiatry, vol. 74, May 2003, pp. 558-560, doi: 10.1136/jnnp.74.5.558.

[71] L.E. Berk. Child development, 7th ed. Boston, MA: Allyn and Bacon, 2006.

[72] S.J. Shettleworth. Behavior, cognition, evolution, 2nd ed. Oxford: Oxford University Press, 2010.

[73] P.M. Gopych, "Generalization by computation through memory," Int. J. Inf. Theo. Appl., vol. 13, Apr.-Jun. 2006, pp. 145-157.

[74] T. Xu, X. Yu, A.J. Perlik, W.F. Tobin, J.A. Zweig, K. Tennant et al., "Rapid formation and selective stabilization of synapses for enduring motor memories," Nature, vol. 462, Dec. 2009, pp. 915-919, doi: 10.1038/nature08389.

[75] G. Yang, F. Pan, and W.-B. Gan, "Stably maintained dendritic spines are associated with lifelong memories," Nature, vol. 462, Dec. 2009, pp. 920-924, doi: 10.1038/nature08557.

[76] N.E. Ziv and E. Ahissar, "New tricks and old spines," Nature, vol. 462, Dec. 2009, pp. 859-861, doi: 10.1038/462859a.

[77] G. Rizzolatti and L. Craighero, "The mirror-neuron system," Ann. Rev. Neurosci., vol. 27, 2004, pp. 169-192, doi: 10.1146/annurev.neuro.27.070203.144230.

[78] C. Keysers and V. Gazzole, "Expanding the mirror: vicarious activity for actions, emotions, and sensations" Curr. Opin. Neurobiol., vol. 19, Dec. 2009, pp. 666-671, doi: 10.1016/j.conb.2009.10.006.

[79] C. Keysers, J.H. Kaas, and V. Gazzole, "Somatosensation in social perception," Nat. Rev. Neurosci., vol. 11, Jun. 2010, pp. 417-428, doi: 10.1038/nrn2833.

[80] A.P. Yonelinas, N.E. Kroll, J.R. Quamme, M.M. Lazzara, M.J. Sauve et al., "Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity," Nat. Neurosci., vol. 5, Oct. 2002, pp. 1236-1241.

[81] P.M. Gopych, "Sensitivity and bias within the binary signal detection theory, BSDT," Int. J. Inf. Theo. Appl., vol. 11 Oct.-Dec. 2004, pp. 318-328.

[82] B. Naundorf, F. Wolf, and M. Volgushev, "Unique features of action potential initiation in cortical neurons," Nature, vol. 440, Apr. 2006, pp. 1060-1063, doi: 10.1038/nature04610.

[83] J. Klüver and C. Klüver. On communication. An interdisciplinary and mathematical approach. Dortrecht, The Netherlands: Springer, 2007.

[84] M.C. Frank and N.D. Goodman, "Predicting pragmatic reasoning in language games," Science, vol. 336, May 2012, pp. 998, doi: 10.1126/science.1218633.

[85] C. Kemp and T. Regier, "Kinship categories across languages reflect general communicative principles," Science, vol. 336, May 2012, pp. 1049-1054, doi: 10.1126/science.1218811.

[86] S. Baldassi, N. Megna, and D. Burr, "Visual clutter causes high-magnitude errors," PloS Biol., vol. 4, 2006, e56.

[87] G. Rhodes and L. Jeffery, "Adaptive norm-based coding of facial identity, " Vision Res., vol. 46, Sep. 2006, pp. 2977–2987, doi: 10.1016/j.visres.2006.03.002.

[88] J. Klüver, "A mathematical theory of communication: meaning, information, and topology," Complexity, vol. 16, Jan.-Feb. 2011, pp. 10-26, doi: 10.1002/cplx.20317.

[89] P.M. Gopych, "ROC curves within the framework of neural network assembly memory model: some analytic results," Int. J. Inf. Theo. Appl., vol. 10, Apr.-Jun. 2003, pp. 189-197.

[90] V. Gallese and A. Goldman, "Mirror-neurons and the simulation theory of mind reading," Trends Cogn. Sci., vol. 2, Dec. 1998, pp. 493-501, doi: 10.1016/S1364-6613(98)01262-5.

[91] M. Iacoboni, "Neural mechanisms of imitation," Curr. Opin. Neurobiol., vol. 15, Dec. 2005, pp. 632-637 doi: 10.1016/j.conb.2005.10.010.

[92] J.M. Kilner, K.J. Friston, and C.D. Frith, "The mirror-neuron system: a Bayesian perspective," Neuroreport, vol. 16, Apr. 2007, pp. 619-623, doi: 10.1097/WNR.0b013e3281139ed0.

[93] M.A. Arbib, "From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics," Behav. Brain Sci., vol. 28, Apr. 2005, pp. 105-167.

[94] M.A. Arbib. How the brain got language: the mirror neural system hypothesis. Cambridge, MA: MIT Press, 2012.

[95] J.B. Bonaiuto, E. Rosta, and M.A. Arbib, "Extending the mirror neuron model, I: audible actions and invisible grasps," Biol. Cybern., vol. 96, Feb. 2007, pp. 9-39, doi: 10.1007/s00422-006-0110-8.

[96] J.B. Bonaiuto and M.A. Arbib, "Extending the mirror neuron model, II: what did I just do? A new role for mirror neurons," Biol. Cybern., vol. 102, Apr. 2010, pp. 341-359, doi: 10.1007/s00422-010-0371-0.

[97] N. Chomsky. The minimalist program. Cambridge, MA: MIT Press, 1997.

[98] W.J.M. Levelt. An introduction to the theory of formal languages and automata. Amsterdam-Philadelphia: John Benjamins Publ., 2008.

[99] J.H. Hopcroft and J.D. Ullman. Formal languages and their relations to automata. Reading, MA: Addison-Wisley, 1969.

[100] R. Carnap. Meaning and necessity. Chicago, Illinois: Chicago University Press, 1948.

[101] B. Russell, "On denoting," in Philosophy for the 21st century, S.M. Cahn, Ed. Oxford: Oxford University Press, 2003, pp. 512-518. Originally published B. Russell, "On denoting," Mind, 1905, vol. 14, pp. 479-493.