Primary Language for Semantic Computations and Communication without Syntax

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

***Abstract*—Recent binary signal detection theory (BSDT), extended by its infinity hypothesis (infinity of common prehistory of universe, life, and mind), is called extended BSDT. Its basic notions underlie BSDT primary language (PL, a hypothetical genuine mathematics used by animals for their internal computations). BSDT PL operates with meaningful words defined as finite binary affixes to infinite binary strings that have common infinite initial parts. In this paper, by an analysis of composite PL words (sentences), it has been demonstrated that meanings of their constituents can only conditionally be related. Meanings of composite words taken as a whole are perceived unambiguously and, under condition that communicated parties have common evolution history (respective infinite strings share their infinite initial part), make possible reliable, in particular non-syntactic, meaningful (semantic) communication. It has also been shown the BSDT neural network learning paradigm, "one-memory-trace-per-one-network", and super-Turing hyper-computations are the mandatory requirements for doing semantic computations and for unambiguous understanding of meanings of finite symbolic messages. Numerical and empirical evidences of some PL predictions and potential PL applications are briefly discussed.**

Keywords-context; infinity; meaning; subjectivity; categorization; complexity; super-Turing computations.

I. INTRODUCTION

Many researchers believe the problem of linguistic meaning cannot adequately be solved without solving the problem of consciousness. Taken together, recent binary signal detection theory, BSDT [1], and its atom of consciousness model, AOCM [2], give a chance of finding a new solution to the problem of meaning. The result is a primary language, PL [3], implementing John von Neumann's 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" [4, p. 92]. BSDT PL [3] is thus a low-level language of symbols (spike patterns) probably used by the nervous system for its internal computations. It may also be used as a precursor to higher-level (including natural) languages that may be built with its help.

In this paper, basic BSDT PL's assumptions (new infinity hypothesis and meaning and subjectivity defined with its help) and some formalism details (computations with infinite binary strings that share infinite initial part) are briefly summarized. Within the PL framework for describing the meanings of components of composite words, the notion of conditional meaning for PL words of different meaning complexity [3] is introduced and on its ground their meaning ambiguity is discussed. For the first time, it will be explained in which way the rigid certainty of meanings of given-level (given meaning complexity) PL words and inevitable ambiguity of relations between meanings of different-level (of different meaning complexity) PL words ensure the richness of PL semantics, and how the PL provides a possibility of reliable *communication without syntax* (or even without any language at all) between animals of the same (or relative) species. For the success, it is needed to fulfill major requirement of BSDT PL infinity hypothesis – for different communicators infinite strings describing the meaning of a finite symbolic message must share their infinite initial part or, in other words, communicators must share significant part of their evo-devo history. Non-syntactic communication and communication without any language at all are the problems 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 [5].

Computational, neuroscience, and psychological evidences of some PL predictions are discussed (the latter becomes possible because there are elements of psychology, i.e., meanings, in the background of PL mathematics). We conclude the network learning paradigm "one-memory-trace-per-one-network" and super-Turing hyper-computations are the mandatory requirements for successful semantic computations and for unambiguous understanding of the meanings of finite symbolic messages. It is claimed that, in living organisms (where meanings of communicated messages are crucially important), Turing and super-Turing computations are the everyday, routine, ubiquitous practice.

The paper consists of Sections I to VII and reference list.

II. BSDT PL INFINITY HYPOTHESIS

Extended BSDT, eBSDT, is the BSDT extended by new *infinity hypothesis* [2, 3] implying the infinity of common prehistory of the universe, life, the mind, language, and society and, according to which, main the eBSDT quantities – meanings of finite-in-length symbolic messages – are defined as infinite symbolic strings that have common "in the past" infinite initial parts. BSDT PL [3] and BSDT AOCM [2] are grounded on the eBSDT and closely related because *meanings* are interpreted as *subjective* experiences of respective feelings (qualia) and vice versa [2].

The leading idea is to 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 creation of this device in the course of evolution from the beginning of the world until now, and *the meaning* of the message under consideration [3]. Under the things of the world, we understand any inanimate objects, animate beings, and any relations between/within them. If in an arbitrary chosen place to split infinite on a semi-axis binary string just introduced into two parts then its finite and infinite fractions could be thought of respectively as *the name* of a thing and *the context* in which this name appears. Such a generalization allows the defining of an infinite but countable number of meaningful finite binary messages (PL words) that would name all the known (and unknown but conceivable) things of the world.

Because of the common co-evolution of all the things of the world meanings of their names are to be described at a given moment by different infinite strings with *common infinite initial parts*. The main feature of these strings is that, having at their ends different finite-in-length fractions, they share their infinite initial part (the text written in it does not matter here). Thanks to this property, the lengths of infinite PL meaning descriptions which differ, if they are indeed different, may explicitly be compared (Section III B). Since the meaning of a name is simultaneously the physical device designed to recognize exactly this name, name's meaning is the property of perceiving organism (sensory agent) and, through it only, of the thing whose name is currently under consideration. The same is also the reason why a name's meaning is simultaneously the animal's respective internal (psychological) state or its current *subjective* "first-person" experience or quale [2, 3].

In order for the AOCM/PL to be able to do *semantic computations* (i.e., to operate explicitly with meaningful strings of infinite length), specific *super-Turing* techniques and the implementation of real-world super-Turing physical devices are required. These are BSDT ASMs (abstract selectional machines [6]), AOCM/PL's building blocks devoted to processing separate meaningful messages or PL words/names. Because of our infinity hypothesis, the ASMs, AOCM, and PL should be based on notions that appeal to an extent to psychology (to meanings of names) and, for this reason, are *beyond* the scope of traditional mathematics.

III. Elements of BSDT PL Formalism

A. Meaningless and Meaningful Words

All the conceivable *meaningful* PL expressions are defined as infinite spinlike (with components ±1) binary strings $c_{xi}x_j^i$ of the same infinite length, $l(c_{xi}x_j^i) = l(c_{x0}) = \aleph_0$ bits ($\aleph_0$, Georg Cantor's aleph) or of the same *meaning complexity* [3]. They constitute (are the members of) an ultimate or proper class S_{cx0} (the set of strings of the length $l(c_{x0})$ that is not a member of any other set [7], $c_{xi}x_j^i \in S_{cx0}$; the term "proper class" may intuitively be interpreted "as an accumulation of objects which must always remain in a state of development" [8, p. 325]). Given-level (Section III B) $c_{xi}x_j^i$ are uniquely specified (marked/labeled) by their right-most fractions, i-bit strings x_j^i (x_j^i is an affix added to the c_{xi}, c_{xi} is common infinite context for all the x_j^i of the length i with different arrangements of their ±1 components; $i = 0, 1, 2, \ldots$ and $j = 1, 2, \ldots, 2^i$). The number of elements (the cardinality) of the fraction of the S_{cx0} that comprises all the $c_{xi}x_j^i$ with x_j^i not longer than i bits is the sum $\sum 2^k = 2^{i+1} - 1$ ($k = 0, 1, \ldots, i$). Consequently, between naturals in their usual order and all the elements of the whole S_{cx0}, $c_{xi}x_j^i$, a one-to-one correspondence can be established. That means the S_{cx0} is *countable* and its cardinality, $|S_{cx0}|$, equals $\aleph_0$. On the other hand, $|S_{cx0}| = 2^{i+1} - 1$ with $i = \aleph_0$; that is, if the 1s that are inessential in this expression are omitted, it will be the famous formula for the size of the Cantor's continuum.

An affix x_j^i may simultaneously be treated either as the ijth i-length binary string, message, computer code/ algorithm, vector in i-dimensional binary space (i-BS), point in the i-BS, element of the set of 2^i points of the i-BS, PL word or PL name. Depending on the current context, these terms will further be used interchangeably.

A word/name x_j^i is the *meaningless* fraction of a meaningful string $c_{xi}x_j^i$; i.e. such a name gets its meaning from its context and from itself, $M(x_j^i) = c_{xi}x_j^i$. Different x_j^i specify all the conceivable strings $c_{xi}x_j^i$ and at the same time represent all the conceivable mathematical expressions as i-length binary strings – that is, they provide complete (non-Gödelian) arithmetization of these expressions by ordinals/naturals (x_j^i may be treated as ordinals/naturals written down in binary notations). For this reason, x_j^i are also the ijth eBSDT Gödel's numbers, $G_{ij}^x = x_j^i$, enumerating themselves and meaningful strings $c_{xi}x_j^i$. The same x_j^i are also the ijth partial Gregory Chaitin's Ω, $\Omega_{ij}^x = x_j^i$ (the ijth halting probabilities [9] for arbitrary binary computer codes not longer than i bits running on the ijth Chaitin's self-delimiting computers or, we hypothesize, on respective BSDT ASMs). x_j^i as well as G_{ij}^x and Ω_{ij}^x are *random* and *incomputable* because they are randomly selected from 2^i different binary i-lengh strings and, then, assigned to things to be named [3]. In other words, x_j^i, G_{ij}^x, and Ω_{ij}^x provide irreducible descriptions (specifications) of these things. The totality of given-level (Section III B) values of x_j^i, Ω_{ij}^x, or G_{ij}^x is the totality of given point-of-view *irreducible* descriptions of all the things of the known world [3].

If string variable x^i consists of variables u^p and v^q then $x^i = u^pv^q$, $i = p + q$; x^i is a string template of i empty cells needed to produce the strings x_j^i by filling these cells in +1s and −1s; u^pv^q is a concatenation of 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 $|S_{vq}| = 2^q$; if $p \le q \le i$, $S_{up} \subseteq S_{vq} \subseteq S_{xi}$. Composite set/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, p and q are the measures of discrete "colored" non-localities of vectors in spaces S_{vq} and S_{up}, respectively [3]. Three-dimensional blue-and-red binary space ("colored Boolean cube") has earlier independently been used for representing the Boolean functions of one-dimensional cell automata, e.g., [10, ch. 6]. The rainbow of colors in finite-dimensional binary spaces here discussed is a direct generalization [3] of the two-color case [11].

B. Categories, Meaningful Words of Different Levels

The form $C(x^i) = c_{xi}x^i$ (it is a concatenation of string c_{xi} and string template x^i) defines a *category* (notion or concept) of meaningful names $c_{xi}x_j^i$. 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 $c_{xi} = c_{up}$, $x_j^i = u_r^p v_s^q$ (i.e., x_j^i is a composite string), $c_{vq} = c_{up}u_r^p$, $x_j^i \in S_{xi}$, $u_r^p \in S_{up}$, and $v_s^q \in S_{vq}$. Infinite strings $c_{up}u_r^p \in S_{cu0}$ and $c_{vq}v_s^q \in S_{cv0}$ are the members of different ultimate classes, S_{cu0} and S_{cv0}, but, because of our infinity hypothesis implying that $c_{xi} = c_{up}$, the lengths of $c_{up}u_r^p$ and $c_{vq}v_s^q$ 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, $S_{cv0} = S_{cx0}$ and $l(c_{xi}x_j^i) = l(c_{vq}v_s^q)$; infinite words [12] of automata theory have no such properties and remain within the framework of traditional mathematics). The form $C(v^q) = (c_{up}u_r^p)v^q = c_{vq}v^q$ with values $(c_{up}u_r^p)v_s^q = c_{vq}v_s^q$ provides a temporal sub-categorization of members of the category $C(x^i) = c_{xi}x^i$. Since, under condition $c_{xi} = c_{up}$, $c_{xi}x_j^i$ and $c_{up}u_r^p$ are of different lengths (have different meaning complexities), we refer to their affixes as names of different *levels*: the level of x_j^i is zero, the level of u_r^p (if it is a fraction of x_j^i) is $q = i - p$, i.e. the number of "ignored" bits that differentiate the length of $c_{up}u_r^p$ from the length of $c_{xi}x_j^i$ (q also defines discrete colored 2^q-state non-locality of u_r^p; as $l(c_{xi}x_j^i) = l(c_{vq}v_s^q)$, v_s^q is also a zero-level name). *Only zero-level names get definite meanings* (see Fig. 1), namely x_j^i in $c_{xi}x_j^i$ or v_s^q in $(c_{up}u_r^p)v_s^q$; the string u_r^p has no definite meaning in $(c_{up}u_r^p)v_s^q$ but it gets a strictly defined meaning as a right-most (zero-level) fraction of $c_{ur}u_r^p$ (if u_r^p is not a fraction of any composite string). A right-most (zero-level) item of a composite meaningful name is called the "*focal*" item (it occupies dynamically created "focus of attention"), a composite name's non-focal item produces a focal name's "*fringe*" (by analogy with fringes of memory and consciousness [2, 3]) or its short-range immediate context.

All the PL's meaningful strings are defined on a semi-axis (i.e. they are "one-side infinite"), have the lengths $\aleph_0$ bits (i.e. they are countable), and must *always* be arranged in a way when they *share their infinite initial part,* the length of which is again $\aleph_0$. Since infinite meaningful strings are arranged in such a way, their beginnings (bit-by-bit common infinite initial part) are always the same but their end-points may not coincide and one of these strings may in general be a number of bits longer or shorter than the other (in other words, they may have larger or smaller meaning complexity [3]). Therefore the strings that are of the same infinite length in the sense of Cantor (that are countable) may be of different infinite length (meaning complexity) in the sense of the BSDT PL. The level of a PL name is the measure of such a difference or the *relative* measure of complexity of meanings; absolute measure (the length of a meaningful string taken separately) is useless for comparing meaning complexities because all meaningful strings taken separately have the same length, $\aleph_0$. Consequently, the notions of a name's meaning complexity and a name's level (relative measure of its meaning complexity) exist in the framework of the BSDT PL only and have their roots in its infinity hypothesis. Meaning complexity embraces given the context Shannon-type ensemble complexity (the length of x_j^i in bits) specifying a name's statistical properties and Kolmogorov-type algorithmic complexity (the length in bits, $\aleph_0$, of computer program, $c_{xi}x_j^i$, that gives complete irreducible infinite description of the ASM that selects the x_j^i) specifying the complexity of devices selecting the names of given ensemble complexity [3].

Composite names $x_j^i = u_r^p v_s^q$, the values of $x^i = u^p v^q$, are thought of as PL *sentences*. If so, the value of a focal string variable, e.g. v_s^q, corresponds to a sentence's feature/ attribute that is currently in the focus of attention; its fringe, e.g. u_r^p, is the fringe of an animal's memory or consciousness. A composite name's "holophrasical" presentation, e.g. x_j^i, corresponds to the perception/ understanding of a sentence as a whole whereas its serial presentation (e.g., a sequence of v_s^q with $1 \le q \le i$) represents the sentence's serial perception/understanding as a sequence of its meaningful fractions or "words". Given the context, different zero-level names are *synonyms* naming the same thing in different ways (e.g., x_j^i is one of 2^i synonyms defined given the context c_{xi} and v_s^q is one of 2^q synonyms defined given the context c_{vq}); if $c_{vq} = c_{xi}u_r^p$ (i.e., if v_s^q is a focal fraction of compound name $x_j^i = u_r^p v_s^q$), both types of synonyms describe the same thing but of different points of view (grounds for the understanding). Any paraphrase of PL sentences (other choice of their "focal" fractions) cannot change their whole meanings and in that sense BSDT PL lacks "compositional semantics" [3].

As composite words are treated as PL sentences, the set of rules defining the relations of meaning of a composite word to meanings of its constituents represent the PL syntax. If internal structure of PL words/sentences is ignored and their meanings are only perceived as a whole then communication with their help do not appeal to PL syntax and, consequently, is carried out *without syntax*.

C. Meanings, Subjective Experiencies (Qualia) and Truths

Given the c_{xi}, each string x_j^i is selected by its BSDT ASM(x_j^i) intentionally designed in the course of evolution and tuned in the course of its individual development exactly for this purpose [6]. An infinite, symbolically-written, complete description of evo-devo prehistory of designing this real-world physical ASM(x_j^i) is the *explicit meaning* of x_j^i, $M_{\mathrm{expl}}(x_j^i) = c_{xi}x_j^i$. The running of ASM($x_j^i$) itself in its real-world physical form is the *implicit meaning* of x_j^i, $M_{\mathrm{impl}}(x_j^i)$, or internal, "mental" or psychological representation of the thing named by the x_j^i [2, 3]. As $M(x_j^i) = M_{\mathrm{expl}}(x_j^i) = M_{\mathrm{impl}}(x_j^i)$, the meaning of x_j^i given c_{xi} is animal's being in a specific psychological state which is a "quale" (subjective "first-person"/private experience or feeling) of this meaning [2, 3]. In particular, the meaning of a category of names is a set of respective qualia.

The name x_j^i is true if 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 adjoined to each other. If there is no such correct correspondence, meaningful name is false. Since the cardinality of S_{cx0}, $|S_{cx0}| = \aleph_0$, is infinite, the number of PL truths is also potentially infinite and, for any meaningful string, its truth value

$T(c_{xi}x_j^i)$ certainly exists (it is either "true" or "false"). Each true meaningful name, e.g. $c_{xi}x_j^i$, names by definition the ith *real-world* thing given to an animal through its ijth psychological state or, in other words, through the activity of physically implemented real-world ASM(x_j^i) [2, 3]. Thus, for meaningful names, 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 which does not hold for PL *meaningful zero-level* names, $c_{xi}x_j^i$. This inference is caused by the fact that PL name meanings are always the ones that animals/humans *actually* keep in mind. It is also the reason why BSDT PL works so well as a primary language (to survive, an animal does not lie to itself) [3].

As truth values $T(c_{xi}x_j^i)$ are never communicated together with x_j^i, they should always be discovered in the process of decoding (*understanding*) the received names x_j^i and confirmed by checking their correspondence to the reality or, more directly, to an animal's respective psychological state. At the same time, a zero-level name's fringe items (names), due to their non-locality, have no meanings but only *conditional meanings* (Section IV and Fig. 1). Hence, for PL names of different levels (meaning complexities) relations between their meanings remain fundamentally ambiguous. This vagueness is a BSDT PL counterpart to Gödel's incompleteness (axioms, theorems, and meta-mathematical expressions for which Gödel's results hold are, in our terms, an infinite fraction of infinite in number *meaningless* strings x_j^i) [2,3].

IV. BSDT PL Meaning Ambiguity

It is assumed that, in a meaningful string $c_{xi}x_j^i$, its context c_{xi} and its name x_j^i describe respectively the *static* part of the ASM(x_j^i) selecting the x_j^i (its "hardware" already fixed in the course of evolution) and the *dynamic* part of the ASM(x_j^i) (its "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 ith thing named by the x_j^i; the jth arrangement of ±1 components of x_j^i is the jth PL description of this ith 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), not the complexity of the thing named by x_j^i.

If $x_j^i = u_r^p v_s^q$, 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 pth and the qth "virtual" things (i.e. of temporally highlighted/allocated fractions of the ith 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 ith thing. Composite names essentially enrich the 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).

Zero-level names x_j^i and v_s^q (v_s^q is a part of $x_j^i = u_r^p v_s^q$) 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, $c_{vq} = c_{up}u_r^p$; Fig. 1(a)). v_s^q is selected by the ASM(v_s^q) that is "virtual" with respect to the ASM(x_j^i), here $c_{xi} = c_{up}$ and for x_j^i and v_s^q their common infinite context is c_{xi}. Thus, ASM(x_j^i) can temporally serve as 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 ($c_{vq}v_s^q$ is simply a variant of $c_{xi}x_j^i$).

If, given the context, $c_{xi} = c_{up}$, names x_j^i and u_r^p are both at the level of zero, then their meanings are to be of *different* proper classes and should have *different* meaning complexities (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 Fig. 1(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 (Fig. 1(a)) and u_r^p (Fig. 1(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 Fig. 1(a)) are required to uniquely establish their correspondence.

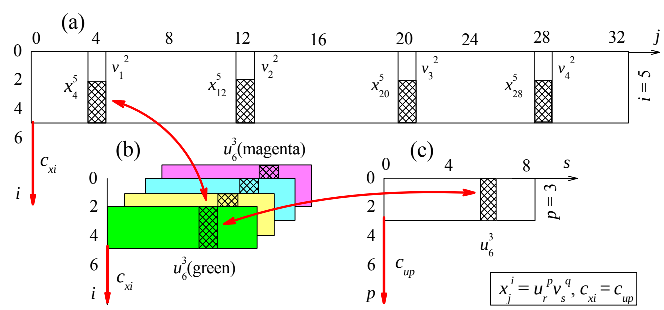

Figure 1. Comparing given the context different-level BSDT PL meaningful names of different proper classes: (a) zero-level names, (b) colored zero-level names corresponding to names in (a), (c) zero-level names that are predecessors to names in (a) and counterparts to names in (b).

If u_r^p is a q-level fringe of zero-level focal string v_s^q and they are both fractions of $x_j^i = u_r^p v_s^q$ (Fig. 1(a)) then u_r^p has *no* meaning (Section III B). But it could get a definite *conditional meaning* if one supposes that u_r^p is conditioned by the color of a zero-level name u_r^p(*color*) selected by a respective q-stages-back-in-evolution ASM (zero-level x_j^i in Fig. 1(a) are unambiguously related to zero-level names u_r^p(*color*) in Fig. 1(b)). u_r^p(*color*) and u_r^p conditioned by one of q colors, though, have conditional but certain meanings. But once colors are deleted (only uncolored strings are used in computations) 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 meaning

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-224-0

of x_j^i and conditional meaning of u_r^p (Fig. 1). The origin of the conditioned relationships just explained between meanings of names of different meaning complexities is the properties of ultimate/proper classes caused in turn by BSDT infinity hypothesis. Of this follows that famous Burali-Forti paradox "there can be two transfinite (ordinal) numbers, *a* and *b*, such that *a* neither equal to, greater than, nor smaller than *b*" [13, p. 157] means in our terms that meanings of PL names, whose meaning complexities differ in *q* bits, can only be compared with 2^q-state uncertainty (Fig. 1).

In Fig. 1, 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* and the width $|S_{xi}| = 2^i = 32$, the *ij*th bar of the height *i* in the *j*th horizontal position designates the name $x_j^i = G_{ij}^x = \Omega_{ij}^x$; bars $x_4^5 = u_6^3 v_1^2$, $x_{12}^5 = u_6^3 v_2^2$, $x_{20}^5 = u_6^3 v_3^2$ and $x_{28}^5 = u_6^3 v_4^2$ that correspond to four colored highlighted bars in (b) are also highlighted (u_6^3 is *q*-level fringe of zero-level v_s^q that is a focal fraction of 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*) = G_r^p(*color*) = Ω_r^p(*color*) corresponding to names x_j^i in (a) (the word *color* is a parameter, names u_r^p(*color*) are selected by conditioned *q*-stages-back-in-evolution ASM and 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 (focal) names u_r^p of meaningful strings $c_{up}u_r^p$ (they name evolutionary predecessors of the thing named by the x_j^i); the *pr*th bar of the height *p* in the *r*th horizontal position (it is shaded) designates $u_r^p = G_{pr}^u = \Omega_{pr}^u$ for the case $u_6^3 = G_{3,6}^u = \Omega_{3,6}^u$. In (a), (b), and (c), strings that are numerically equivalent to the u_6^3 are shaded in the same way; contexts (they are shown as bold arrows) are equal to each other bit by bit, $c_{xi} = c_{up}$. Uncolored and colored names name real-world and conditioned ("virtual") things, respectively. A bijection, $x_j^i \leftrightarrow u_r^p$(*color*), exists between names in (a) and names in (b); it may be e.g. $x_{28}^5 \leftrightarrow u_6^3$(*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), e.g. $u_r^p \leftrightarrow u_r^p$(*green*) (once it is established, other conceivable bijections, e.g. $u_r^p \leftrightarrow u_r^p$(*magenta*), become impossible). If 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). 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.

V. NUMERICAL AND EMPIRICAL BSDT PL VALIDATIONS

Semantic computations produce meaningful results of meaningful data. BSDT PL computations are exactly of this type because we imply that the meaning (infinite context) of any finite mathematical expression is always taken into account when any formal operations (computations) are being done on it. For this reason, elements of psychology (meanings) are always involved in semantic computations and their completely formal (i.e. independent on meanings) descriptions become strictly speaking impossible. Thanks to this fact it does become possible to verify the methods of proposed PL mathematics by methods of psychology and neuroscience or, in other words, by comparing PL computations with internal computational mechanisms that are actually in live animals/humans.

A. Solving Communication Paradox

In Section III B and Section IV we saw that only zero-level names whose internal structure (the manifold of their possible focal and fringe constituents) is ignored or, in other words, only those PL sentences that are presented *without syntax* and perceived "holophrasically" have given context unambiguous meanings. This fact and the fact that meanings of PL meaningful names are the ones that animals/humans *actually* keep in mind [3] make the BSDT PL an appropriate tool for the description of communication without syntax (or without any language at all) that is typical for animals and human infants, e.g. [5] and references therein. We hypothesize: communication without syntax (it is exhibited as an animal's basic/inherent behaviors that truly reflect its respective inner states or behavioristic part of animal's cognition) suffices to support *the simplest* animal sociality.

Since complete meaningful descriptions of PL names, $c_{xi}x_j^i$, are fundamentally *infinite*, during any finite time period they can never be communicated in full even in principle while in fact many times a day everybody observes in others and experiences him/herself successful meaningful information exchanges. This *communication paradox* [2, 3] speaks of everyday, routine, ubiquitous use of super-Turing computations in human meaningful and socially-important communication. The communication paradox can be solved by appealing to BSDT infinity hypothesis [2, 3] and the technique of BSDT ASMs that are super-Turing devices with programmatic and computational processes that are completely separate in time (ASMs do not waste their computational resources on serving themselves and, for this reason, are faster than *universal* Turing machines [6]).

But, dividing the programming and program running is insufficient to overcome the communication paradox. To cope with it, let us additionally assume that the ASM-transmitter and the ASM-receiver share in full their prehistory, i.e., let 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 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 meaningful information exchange, the transmitter and the receiver are to be exactly physically, structurally, and functionally equivalent (are to be "mirror" replicas or "clones" of each other).

 ISBN: 978-1-61208-224-0

Several important PL predictions come out.

B. Coding by Synaptic Assemblies

Where meanings are essential (e.g., in living organisms) BSDT network learning paradigm “one-memory-trace-per-one-network” [14] 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 memory traces as possible but it is well supported by recent empirical neuroscience finding of coding by synaptic assemblies [15, 16]. In these experiments, mice were trained to perform new motor tasks and in living animals changes in the number of synaptic contacts associated with learning new skills were measured. In complete accordance with the BSDT assumption [14] that each new memory trace should be written down in an always new separate network (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 [17, p. 859].

C. Super-Turing Computations by Mirror ASMs

To ensure correct understanding of meanings of finite symbolic communicated messages, the ASM-transmitter and ASM-receiver that are the mirror replicas of each other need to be used. 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 Turing-type computations, which have been programmed beforehand, with finite-length strings x_j^i, e.g., as in [14]. Mirror ASMs *physically* divide infinite meaningful message to be processed into infinite, c_{xi}, and finite, x_j^i, parts and take the former into account as their exactly identical “hardware” and “software”, 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. 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” that fix the common infinitely long context needed to finish meaningful super-Turing computations of current interest over a finite time period. 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. [18, 19] and numerous references therein. The ASM/mirror-ASM computational system just described and the neuron/mirror-neuron circuitries already observed [18, 19] may respectively be treated as theoretical and real-world implementations of super-Turing machines with infinite inputs, which until now have been hypothetical, e.g. [20], that are to be capable of computing with infinite strings or, what is the same, with real-valued/continuous quantities.

D. Knowing Memory Performance without Knowing Memory Record

Since BSDT PL employs a non-Gödelian (but envisaged by Gödel [21]) arithmetization by ordinals/naturals x_j^i and since these ordinals/naturals are given context *randomly* chosen to name the things to be named, the following effect has been predicted [3]. By examining in an experiment an ASM hierarchy (neural subspace [14]) that generates the meaning of a trace x_j^i, all the parameters describing the ASM selecting the x_j^i may successfully be found but the content of x_j^i – specific 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 cannot empirically be found. This rather surprising prediction has been corroborated well by numerical BSDT analysis [22] of receiver operating characteristics (ROCs, functions providing memory performance) measured in groups of brain patients and control healthy subjects. In such a way the idea of non-Gödelian BSDT arithmetization of meaningful mathematical expressions by ordinals/naturals *randomly* chosen given the context of a set of ordinals/naturals with their given upper limit has numerically and empirically been substantiated.

VI. Examples and Potential PL Applications

Any formal axiomatic system (FAS, it comprises all its axioms and theorems) is supposed to represent an *infinite fraction* of meaningless finite binary strings x_j^i related to particular proper class, S_{cx0} (Section III A). For this reason, FAS computations are also PL computations and numerous available computational results e.g. in physics or biology may be treated as their examples performed given a context defined formally and informally. A separate infinite PL string that gives a meaning to a finite symbolic message x_j^i (e.g., a formula written in binary notations) includes descriptions of the FAS formalism needed to derive it and of the problem that gives it physical sense. This picture is another representation of formal and informal knowledge from e.g. a book and gives nothing new, except of drawing attention to the fact that manipulations with numbers are meaningless until a giving-the-meaning context is added.

The situation changes dramatically once one wants to communicate this formula’s *meaning* to someone else. Let us consider a lecturer in a lecture room. In the beginning, he and his students have different knowledge on the formula of interest and students cannot 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 of understanding this formula/message. At the end of the lecture, for the lecturer and for his students, infinite PL strings describing 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 [2]. If it is not, a misunderstanding arises. How, for members of a social (semantic) network, the difference in their previous knowledge influences on understanding meaningful messages may empirically be estimated as described in [23].

In this example, non-syntactic messages represent a very small fraction of general flow of information. Among them it may e.g. be the fact that the lecturer is walking when he gives his 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 and have the same *innate* bodily infrastructure (in particular mirror neuron system) to produce and perceive/understand walking. Humans/animals produce and perceive such message automatically with practically no chance of misunderstanding because its meaning is provided, for all of them, by *innate* infinite PL strings of the same length that are equivalent "in the past".

The range of potential PL applications covers everything where meanings are important. If they become inessential, the PL can be reduced to traditional mathematics (a FAS).

VII. Conclusions

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 discrete computational part of intelligence of animals of poor sociality and, consequently, to design the discrete computational part of intelligence of artificial devices (e.g., robots) or computer codes mimicking the behavior of such animals. At the same time, the PL is unable to explain the mechanism of dividing its names (sentences) into focal and fringe components and, consequently, of directing an animal's attention to particular thing – we hope it may be done by methods beyond the BSDT. 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.

In contrast to formal languages that are in end the products of a *finitely* defined calculus, BSDT PL is a calculus of finite binary strings (spike patterns or "symbols") with *infinitely* defined contexts. It is grounded on 1) the BSDT [1] providing the technique of encoding/decoding in binary finite-dimensional spaces (BSDT ASMs [6] implement PL's inference rules) and 2) the new infinity hypothesis [2, 3] providing the technique of super-Turing (semantic) computations with infinite binary strings that share their infinite initial part. 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 different-level BSDT PL names that have been established as their fundamental property raises many intriguing problems to be solved in the future.

References

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

[2] 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, G. Ellinas, Eds. Berlin-Heidelberg: Springer, 2009, pp. 54-64, doi:10.1007/978-3-642-04277-5_6.

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

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

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

[6] 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, X. Yao, Eds. Berlin-Heidelberg: Springer, 2007, pp. 198-208, doi:10.1007/978-3-540-77226-2_21.

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

[8] 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.

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

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

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

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

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

[14] 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.

[15] 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.

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

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

[18] 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.

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

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

[21] 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.

[22] 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.

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