I. Egorychev

CATEGORIAL ANALYSIS OF A. KURPATOV'S "METHODOLOGY OF THOUGHT" IN CONTEXT OF PERSPECTIVE AGI[1] DEVELOPEMENT

"When we establish a considered classification,
when we say that a cat and a dog resembe each other
less then two greyhounds do even if both are tame
or embalmed, even if both are frenzied,
even if both have just broken the water pitcher,
what is the ground on which we are able to establish the validity
of this classification with complete certainty?"

Foucault M. The Order of Things.

Abstract. In the article we apply language of category theory in order to formalize core methodological principles that structure the methodology of thought elaborated by Russian modern psychiatrist and philosopher A. Kurpatov. According to the author such formalization could be useful both from the standpoint of unification of ways of thinking about brain functioning and reasoning in particular, and from the standpoint of search of uniform language of scientific thought in general.

Preliminary remarks

The task we are eager to solve within the limits of declared theme of present research can be formulated as follows: to mathematize core principles of brain's work. We will proceed from the assumption that human's brain is extremely effective machine capable of decision making in situations with

[1] Artificial General Intelligence

radical incompleteness of information. This is so by very the same brain's construction which has been justified and is still justifying itself in a process of natural selection.

It doesn't mean however that given evolutionary construction at all flawless. But provided we clearly understand the genesis of these flaws they could be seen as certain specificity in work of this instrument which just needs to be taken into consideration. Further still, this specificity emerged by no means randomly, but rather as one of the possible nature's solutions of the problem of *speed* in decision making, which was solved precisely by means of sacrifice of *precision* (truth) - quite a legitimate move, statistically speaking.

Therefore, when trying to build general artificial intelligence it would be productive not to emulate human's central nervous system, but rather to reproduce those abstract principles that are implemented in brain's cognitive activity, i. e. its ways of information processing, deduction or making decisions considered as neutral with regard and neurobiological characteristics of living tissue.

At the time, Andrey Kurpatov, eminent Russuian psychotherapist and philosopher, carried out immense scientific work describing patterns of human thought in highly abstract terms. His early theoretical writings - 'Philosophy of psychology' and, especially, 'Tractatus psychosophicus' - are written exactly in this way. In these investigations he uses (as he calls it) "referentless"conceptual apparatus which grasps brain functioning not as a system of specific nerve tracts, but rather as a matrix of regularities, sort of logic of occurring processes.

Earlier we already tried to "translate"basic ideas of Kurpatov's "Tractatus" into the category theory language[2] in order to experiment with the appropriateness and usefulness of categorial formalism in interdisciplinary research

[2] *Егорычев И. Э.* Категорный анализ текста А. В. Курпатова "Tractatus psychosophicus"в контексте перспективных разработок AGI.

related to AGI development. Since this work turned out to be quite satisfactory, so in present research we address to yet another one of Kurpatov's latest theoretical texts called "Methodology of Thought. A Draft where specific character of psyche's action once again expressed *referentlessly* - as some kind of "informal logic of processes".

"In order to express this logic, - Kurpatov writes – we will need following invariants: "intellectual function", which describes all possible operations in the given system, "intellectual object", by which we have in mind any particular wholeness that can be singled out in the space, as well as "entity", which we need exactly for uncovering these particular "intellectual objects".[3]

So, as distinct from "principles" of "Philosophy of Psychology", "things" and their "relations"of "Tractatus Psychosophicus in this present case we will have to formalize quite different set of concepts ("invariants"): "intellectual object", "intellectual function", "entity" and, as we shall see below - "invariant" itself.

Let us start with a little deeper plunging in the context: Kurpatov insists that methodology of knowledge (scientific theory) could be efficacious only if a) it is non-referential; and b) all its non-referential elements are defined through each other.

Although I do not believe that this task could be accomplished entirely as long as even such highly formalized sciences as mathematics and physics are forced to operate with so called primitive, or indefinable notions[4] (number, set, point, mass, time etc.), theory's usage of maximally general notions with minimum of content is totally legitimate - we witness it with the plenty of evidence during all history of development of human thought as such. But to have a precise grasp on how a certain notion works within particular theory we definitely need a context.

[3] *Kurpatov A.* "Thought. Systematic investigation". - SPb, 2019. P.131
[4] `https://en.wikipedia.org/wiki/Primitive_notion`

Following Kurpatov, in our investigation we want, on the one hand, to rely on scientific results of present day neurophysiology, and try not to contradict them in those hypotheses which can not be confirmed at the moment. On the other hand, we want more - we want to handle human thought as "a process of manipulating with intellectual objects while sought by very same neurophysiology one-to-one correspondence between the set of states of neurons with the set of so called mental content (which later we will call "world of intellectual function) is far from being established. Moreover, this task may never be accomplished. From this being said, however, it does not follow that we tend to consider mind as some sort of independed substance, tabooed for any kind of systematic research. Allow me to quote "Methodology of Thought"once again (as readers may guess we will address to this text quite often): "Modern physiology looks a lot like invention of telescope in astronomy or invention of microscope in biology: it investigates organ that *produces* mentality that, in its turn, produces consciousness".[5] (*My italics.* - I. E.). In other words, from the very beginning we deliberately place ourselves on a footing of some fixed materialistic theory of consciousness, or *cybernetic reductionism*, if you will. From such point of view all manifest "phenomenology" is interpreted as *user illusion*, while consciousness itself is considered a virtual machine installed on material substrate that has been "shaped"by evolutionary algorithm in a very special way. Computer metaphor we use seems even more relevant in case we set a goal to apply Kurpatov's "invariants"in computer science. Let's see in more detail what does it give us.

By *virtual machine* computer scientists normally understand a software (list of instructions) installed on some real machine B (on a computer - an actual hardware made of silicon chips, capacitors and wires and so forth) in order to perfectly emulate another real machine A. As philosopher Daniel Dennett once put it: "A virtual machine is what you get when you impose

[5] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 28

a particular pattern of instructions (more literally, dispositions) on a real machine that has lots of plasticity—interacting parts that can be in many different states."[6] For example, on Macintosh computer we can use software programs that run only on Mac OS operating system. And if you have a valuable software program written for IBM PC it won't run on your Mac. Instead one can install on her Mac specially developed software (Parallels Desktop, for one) which will replicate IBM physical architecture, on which appropriate operating system (Windows) can be installed, on which any programs written for "real IMB"will run.

In more extended sense we could call virtual machine both operating system itself and any software program at all which turns our hardware computer into calculator, or CD player, or God knows what else. All those programs don't imitate any actual hardware but create another virtual machine that obeys certain rules, accepts certain inputs etc.

As a result, the main advantage of such extended VMM (virtual machine metaphor) lays in very important option to repeat this "imitation game"with visualization several times. Here is another clarifying example from computer programming: after famous Java Virtual Machine (JVM) were created, software developers got a chance to write so called Java applets - small,integrated directly into the web page, software applications which considerably diversified interactive repertoire of internet websites. These applets includes crosswords, online games, maps, charts and so forth and all of them are written on Java programming language who's source code would be executed on appropriate JVM regardless of which operating system your hardware's running on - appropriate JVM either already is or would be automatically downloaded once you visit certain web page. Sun Microsystems when entering the market with it's revolutionary product even came up with the special slogan WORA (Write Once, Run Anywhere) to emphasize exactly this extremely useful

[6]Dennett, Daniel C.. Intuition Pumps And Other Tools for Thinking (p. 134).

aspect of virtualization.

Now, let's imagine online chess game that is written in Java code that is executed by Java Virtual Machine that is running on Mac OS operating system that is installed on your MacBook. The whole construction can be viewed as machine MacBook pretending to be a MacOS machine pretending to be a Java virtual machine pretending to be Java syntactic machine pretending to be a chess machine.

In other words we have four virtual machines and one real machine that successively imitate each other. What in important here is that while moving chess figures with the cursor we fall a victim of multilevel illusion - user illusion. Because on hardware level there are no any "chess figures', "cursors', "programming languages"or even "strings of ones and zeros"that compose "words"of machine code - there IS only a sequence of alternating with mind blogging speed states of those silicon semiconductors and a flow of electric charge running in a particular fashion through metal wires and conductive layers inside microprocessor.

From such "computer metaphoric"point of view, the main difference between mental organization of human and mentality of any other animal is that former arises as a product of cultural evolution which allows for "installation"on human brain of a huge number of words and other thinking tools (information structures). These information structures eventually form cognitive architecture (top down) radically different from cognitive structure of animals (bottom up). By supplying human minds with the system of *redundant* multiple representations this architecture equips us (humans) with the privileged perspective from which we get limited and biased access to machinery of our brain (of our whole body, to be more precise). Results of these brain workings we involuntary mistakenly attribute either to properties of external world (colors, odors, sounds...) or to internal personal dispositions and responses (expectations, desires, emotions...) All those "familiar things"constitute "world

of intellectual function "scheme of world"and complimentary 'scheme of me' - if we use "Tractatus Phychosophicus"conceptual apparatus[7], or "user interface' - if we use just introduced computer metaphor.

User interface of any software application exists in order to make functions of this application available to users (people) who don't and don't have to know all technical details of how their computers work. User illusion stored in our brains exists for the same reason - it makes our functions available *to some extent* to users (other people) who don't and don't have to know all intricate details of how our brains work. But the same applies and to us: when we get to use it ourselves our position is not privileged - we are the same guest users in our own brains as anybody else.

Italicized "to some extent"is not an accident: it is extremely important. Since from the above it follows that need in clear cut representation of someone's behavior, from evolutionary standpoint, presupposed need in communication, we must emphasize: Darwinian creature that communicate all her states and/or intentions to all hearers will be soon extinct. That's why evolution of communication must be seen as strategic communication which first of all is not entirely linguistic and second - grounded in deception and manipulation rather than purely cooperative behavior. Thus, what must evolve is a *gap*, necessarily emerging as a result of such controlled communication - a buffer, where actual goals and intentions could be doctored enough to be safe for broadcasting. This particular buffer, we must assume, formed the basis for "private workspace"that we now call consciousness.

And final very important point. When viewing consciousness as a virtual machine composed of virtual machines composed of virtual machines etc installed on brain's neural substrate we must understand at the same time that unlike in the case with computer monitor facing end user precisely the latter is final beneficiary/victim of user illusion. In the model where

[7] *Курпатов А.* Психософический трактат.М. 2007. P. 37

ourselves are end users (our brains, to be precise) *the instance of monitor is superfluous*. We will come back to this important issue later, but for now let us underline once again: while scrupulous understanding of neural interactions on molecular level is important scientific goal our present investigation mainly will consist in modeling brain's information processing and converting it in knowledge.

I. Intellectual object

"Human beings have only a weak ability
to process logic, but a very deep core capability of
recognizing patterns."
Kurzweil R. How to Create a Mind

As any other living thing we, humans, preoccupied with self preservation. Strongly generalizing one can say that self preservation directly depends on anticipation, hence will be the more effective the more competitive future will generate given self preserving system. as highly evolved creatures for these purposes we use highly evolved apparatus (that allows us not only to duck in time after someone threw a stone at us, but also to short our stocks in time, after we more or less accurately predicted a decline in oil futures prices) which we conventionally will call "mental apparatus". We say "conventionally"keeping in mind that there are more then a few schools of thought on that matter. As for us, by "mental"we,standing on materialistic position of sorts, understand an *aspect* which is worth distinguishing in neural activity (or rather, *effect* that is produced by such activity), i. e. virtual machine, implemented during evolutionary process on material medium that met with certain requirements. By "certain requirements"we understand the following: just as the wheel being a piece of technology *requires* preexistence of certain specifically developed surfaces (rails, paved roads for its implementation, so the virtual machine of human mind can be deployed only in such "ecosystem where besides language and social interaction writing plays a major role, as well as charting, mapping, diagramming and other endowers of potential intelligence[8], so they could be used to unload many auxiliary virtual machines into the very

[8] P.99

same ecosystem. These virtual machines then can be downloaded from the ecosystem to the individual brain by demand through ability to learn, which presupposes considerable prototypical plasticity of the brain, which in turn is another requirement that needs to be met by this "real machine".

Wheel, scissors, number, long division, alphabet, justice, French language, music, color coding - all of them are intellectual objects, bearers of potential intellect, virtual machines, applets, who's source code can be executed by virtual machine of our mind. But on the lowest levels of virtualization each of these objects is just some "primitive"specific complex of neural interconnections capable of initiating certain behavior (response): *difference that makes a difference*. Actually, each one of these complexes is exactly a one *thing* - of those the world presents itself by. [9] As Kurpatov writes himself: "In effect, this "object of thought"is nothing but separate cortical columns and reflex arcs, connected ("consolidated") as single neural complexes."[10]

This view on "units of thought"agrees very nicely indeed with the results of modern neurobiology. So called *columnar organization* of the cerebral cortex for the first time was discovered and characterized by American neurophysiologist Vernon Benjamin Mountcastle in the 1950s. He observed remarkably unvarying, repeated structure of neocortex and hypothesized that it was composed of cortical columns as its basic units.[11] There are about a half million cortical columns in a human neocortex, each occupying a space about two millimeters high and a half millimeter wide and containing about 60 000 neurons. Later research showed that each column also consists of repeated functional fragments of nervous tissue (about a hundred neurons) united by similar task - to encode and recognize some relatively primitive pattern, or form. That's why now strucrural unit of neocortex defined as said minimal

[9]Proposition 1 of "Tractatus Phychosophicus".
[10]*Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 72
[11]V. B. Mountcastle, "An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System" (1978).

pattern recognizer, while Mountcaste column which is just an aggregate of pattern recognizers called either hyper- or macrocolumn [12]. So, there are on the order of 30 billion neurons or about 300 million pattern recognizers in total in the neocortex.

Kurpatov suggests to consider any intellectual objects as *always-already-derived* from other, more simple intellectual objects: "It would be a mistake to assume presence of some initial (primitime, elementary) intellectual objects. First of all, we must understand that every intellectual object, even the most "simple"one, aggregates from different, separate stimuli (which affect different, separate receptors). Second of all, all these intellectual objects get their appropriate status (disposition, weight, value, sounding) - that of 'intellectual object no sooner than at the moment we endow this intellectual object with some "essence' - in other words, perceive it in some kind of relation with ourselves (where "our-self"considered as any mental content whatsoever), as a "thing"that has certain "meaning-for-me".[13]

Here we must pay attention to several important points at once: first, intellectual object's relation "with me"does not presuppose any awareness of it, or it's representation in consciousness - it should be enough if something would be perceived, differentiated in extent sufficient for this "something"would be taken into account somehow, would make a difference in the future. But, of course, composite intellectual objects may well be represented in consciousness. Second, because, as we mentioned earlier, the main task of cognition is forecasting, or fabrication of the competitive ("farther-sighted") future, the overall presentation of perceived intellectual object will depend on our predisposition,or, as phenomenologists would say - on our *intentionality*[14].

[12] `https://en.wikipedia.org/wiki/Cortical_column`

[13] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 38

[14] Intentionality, in phenomenology, the characteristic of consciousness whereby it is conscious of something i.e., its directedness toward an object.
`https://www.britannica.com/topic/intentionality-philosophy`
"Experiences are intentional. This being-directed-toward is not just joined to the

To put it differently - *expectation affects perception*. Something very similar we find in "Tractatus Psychosophicus": "I do not determine a meaning of the thing deliberately - it is determined by what thing [meaning] is I to myself. Hence: meaning of myself [thing] is always some need that forces me [meaning] to develop a certain thing [meaning] from the World"[15] And third: the very process of thought, for Kurpatov, is not a mind movement, not a "successive transition"from one set of propositions to another (whatever it means), but rather a construction of suitable intellectual object.

Ray Kurzweil, who we quoted in the epigraph to this chapter, and who, while working in Google for more than eight years by now, consistently and successfully uses his own theory of mind which based on hierarchical pattern recognition, expresses a surprisingly similar thought: "To do logical thinking, we need to use the neocortex, which is basically a large pattern recognizer. It is not an ideal mechanism for performing logical transformations, but it is the only facility we have for the job."[16]

Describing the differences between the algorithm of the computer program that defeated World chess champion Garry Kasparov in 1997 and the way of playing chess of the Grand Master, Kurzweil notes that DeepBlue was capable of analyzing the logical implications of 200 million board positions per second. Kasparov, on the other hand, when asked how many positions he could analyze each second, said that it was less than one. But Kasparov new and remembered about 100 000 board positions and compared the situation that he sees at the moment to all 100, 000 board situations that he has mastered, and he does all 100 000 comparisons *simultaneously*[17]. A

experience by way of a mere addition, and occasionally as an accidental reaction, as if experiences could be what they are without the intentional relation. With the intentionality of the experiences there announces itself, rather, the essential structure of the purely psychical". - Edmund Husserl

[15] *Kurpatov A.* Tractatus Psychosophicus.. - M, 2007. P. 36

[16] *Kurzweil, R.* How to Create a Mind. Penguin Publishing Group. London. P.96

[17] There really is a scientific consensus on this point: all of our neurons are processing the

typical medical specialist makes about the same number of operations of pattern comparison/differentiation when she diagnose a patient - he compares discovered symptoms of particular patient with about 100 000 learned characters of illness. Generally speaking, any professional, when solving any intellectual or creative problem, does the same - in order to become a true specialist (to learn this many chunks of knowledge) she needs those notorious 10 000 hours of deliberate practice in chosen domain.

Thus, we see that both Kurpatov and Kurzweil strip the cognition of traditionally attributed dynamics - it is more like "pending"in uncertainty which is almost literally abolished (in Gegel's sense of the German word *Aufheben*) as soon as suitable pattern has been recognized (suitable intellectual object has been constructed).

Taking into account everything mentioned above it is fair to say that even the most elementary intellectual object is indeed not that elemental and really is derived from at least three "elements:

- a set of data (A);
- myself as "the world that presented itself to me as me"(Ω);
- relation with me, which is, generaly speaking, is the function of my internal state/expectation (f).

From now on, every time we will use the term "intellectual object we will keep in mind the following composite construction:

$$\begin{array}{c} A \\ \Big\downarrow f \\ \Omega \end{array}$$

Speaking formally, A and Ω - are "objects"of some sort, and an "arrow" $\longrightarrow$

patterns at the same time.

- is a functional relation. We can see that the whole construction, although it consists of three so called self-sufficient objects, at the same time appears as a certain integrity: relation is always given by some rule, but at the same time the *source* and the *target* of the rule are also constitutive for the relation itself. That's why the whole diagram can be regarded in several aspects:

- with an emphasis on the source object where the arrow departs from - as an illustration of the fact that object A, as part of the relation, is never given to me "in itself"and comprise data that also are not given but rather taken.

- with an emphasis on the target object where the arrow arrives to - in order to make clear that perceived essentially depends on the perceiver. In case an arrow ends in different subject of experience it would be another relation and completely another object;

- with an emphasis on the arrow - in order to show that relation is essentially a process;

- finally, with an emphasis on the whole diagram as on new, self-sufficient object of some sort that was formed by two halves of relation.

Sometimes, to make more distinct that it is the relation that is constitutive for the whole construction we will write: $f : A \longrightarrow \Omega$.

Data that comprise the set A also , generally speaking, are intellectual objects. Collection of data which belong to the set A as its elements also, generally speaking, are composed intellectual objects. But we must start somewhere, so we will proceed from assumingly "primitive" elements $x, y \in A$, keeping in mind that their "primitiveness" is merely an assumption that strongly depend on context.

We must pay special attention to the set Ω as well,because in our construction it will play very specific role - that of the instance of experience (in a sense of affectation). In this subject of experience we will be interested most of all

in his ability of differentiation. Thus, the structure of object Ω which models subject's ability of differentiation must satisfy certain requirements: first of all, it must be a partially ordered set, so to its elements could be assigned more or less high values. That is, a candidate set for Ω must have a partial order structure.[18] In other words, we will use Ω as an existential scale - or simply a ruler, that will measure the differences.[19]

Now let's try to understand more clear what kind of differences elements of set A can have and how we can measure these differences. Referring to the Kurpatov's text once again: "Our mind leans irresistibly to sum up all collection of stimuli into one comprehensible, lucid and supposedly consistent view on reality, i. e. to create a "reality effect". This representation of reality is in fact some sort of filter-interpreter - every new stimulus, when caught by, figuratively speaking, gravity field of certain representation system, inevitably changes its trajectory, so to speak - some of them are repulsed (are ignored), some other, complimentary, are attracted, while some other still are modified (interpreted) in order to fit in existent viewpoint."[20]

As we already mentioned earlier all said above grasps very well with a concept of *expectation*: on every level of perception - from the most primitive, genetically determined[21] ability for differentiation to utterly conscious, abstract concept - we in fact deal with situation, with some expected state of affairs. Respectively, we expect from the elements of the whole situation that they

[18] Every interaction (relation), although it has infinite number of modalities, always can be characterised in terms of intensities of these modalities. And these intensities are partially ordered. See *Egorychev I.* "Categorial analysis of "Tractatus Psychosophicus"

[19] With the values of this ruler we as well might want to do set theoretic operations of union and intersection - it will endow our set of existential values with some basic "logic". See *Egorychev I.* Comments to categorial analysis of "Tractatus Psychosophicus"

[20] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 57

[21] "If we look into the studies on development of face (muzzle, beak) recognition skill which forms in early infancy almost in all animals with relativity large brain we will see that brain tends to develop strong habits of perception based on some sort of instinctive, genetically predisposed preferences." *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 73

will appear in it as this or that. As a result, it makes sense to say to what extent every element included into intellectual object A is, first - different from itself (in a sense of what we expect to see in it's place) and second - to what extent this particular element is relevant to the whole situation, i. e. how *close* is it to the other elements that were discerned in this situation.

Hereinafter we will say that there is a *function of expectation* $\mathrm{Exp}_A : A \times A \to \Omega$ determined on set A which for every two elements $x, y \in A$ assigns a degree of their coherence (proximity) q on existential partially ordered scale Ω so that $\forall a, b, c \in A$:

$$\mathrm{Exp}_A(a, b) = \mathrm{Exp}_A(b, a)$$

$$\mathrm{Exp}_A(a, b) \wedge \mathrm{Exp}_A(b, c) \leq \mathrm{Exp}_A(a, c)$$

We see that these two imposed conditions are quite weak and they become even more comprehensible when we try to interpret the degree of difference (or - proximity) topologically - as proximity in space - and we will measure it as a distance. then our conditions will take a form of axioms of metric[22] with the only difference that proximity of intellectual object to itself don't have to be maximal. Degree of coherence of intellectual object with itself can be understood as degree of proximity to it's own essence (or participation in Idea in Platonic sense) and denoted as $\mathrm{Ess}_A(x)$ We will talk more about essences below.

Weakness of axioms of such generalized metric turns out to be extremely convenient also because these axioms are satisfied by whole class of functions and variability on them can easily be interpreted as variation of internal state of the subject of experience.

Moreover, we see as well that all those degrees will have necessarily local character: the same element $x \in A$ can be measured as more or less coherent

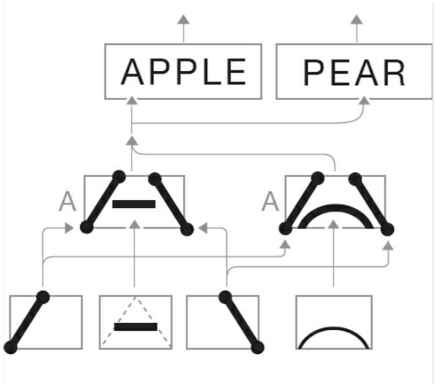
[22] `https://en.wikipedia.org/wiki/Metric_(mathematics)`

depending on nature of the context it is being experienced in, on expectations being laying down (mostly unconsciously) on the situation as a whole[23], as well as on internal state of the subject of experience.

Several examples are to follow which, from our standpoint, could significantly clear up suggested formalism.

First example we will borrow again from Ray Kurzweil's book "How to create mind". Let us remind to all that in accordance with his pattern recognition theory of mind (PRTM) human neocortex is intricately organized hierarchy of modules each one of which on it's own level is responsible for the recognition of patterns of certain type - from, let's say, inclined bar in some graphic character to such highly abstract concepts as irony or justice. It is important to to keep in mind that this hierarchy precisely is the hierarchy of concepts (of intellectual objects) rather then hierarchy of neural complexes that comprise the recognition module itself, because on hardware level the construction of neocortex is only one pattern recognizer high. That's why said hierarchical structure forms not as a result of physical building up of one module above another, but as complex network of neural interconnections which develops in one's ontogenesis. Let's look on some standard fragment of network hierarchy of pattern recognizer of written text which consists of three levels:

[23]such unconscious expectations often called beliefs, i. e. certain set of basic convictions, of faith in that all objects have weight, inability of humans levitate without technical support of some sorts etc. In his "Tractatus Psychosophicus"Andrey Kurpatov called them "outposts of faith".



Elements of letters, whole letters, words - all these are patterns recognized by a module of corresponding level. Each pattern recognizer has number of inputs, a "processor"for recognizing of symbols and unique output. Once lower level modules have recognized their "objects"and sent appropriate signals to the above "floor module of this floor also sends signal that the object it was "responsible"of just has been recognized. Different signals are assigned with different "weights"and some other parameters - such as size of the input signal, as well as variability in this size. If we take phonemes as example of patterns to be recognized then when perceiving someone's articulation of two words "steep"and "step"the expected size of vowel sound "E"in first case will be bigger then in second, as well as expected variability of this size. because normally sound "E"in word "steep"considered as a long vowel although different people may articulate it differently. Sound "E"in word "step"will have smaller expected size, but also a large degree of variability as soon as it still can be pronounced as a long sound. However, we can't say the same about consonant sounds "S"and "T"in the same two words - they will have both small expected sizes and small variability of that size.

We see that "size"in this case means duration of sound, but many other, much more abstract quantities can be coded in the same dimension - such

as sympathy, pleasure, grief etc.[24].As Kurzweil notes himself: "We can draw similarities across rather diverse continuums, as Darwin did when he related the physical *size* of geological canyons to the *amount of differentiation* among species."(My italics. - *I. E.*)[25]

The key point of this example is that module's inputs of a given level receive signals not only from the lower levels, but also from modules positioned higher in hierarchy. And what is more important that input of the latter in summary response could be even more significant. significant. If we are reading from left to right, for example, the word "evolution"and have already recognized all the letters except the last one, the "evolution"recognizer will predict that it is likely to see an "N" in the next position and will send appropriate signal down to the "N"recognizer "saying"that probability to encounter letter "N"is very high. So"N"recognizer adjusts its threshold such that it is more likely to recognize an "N" even if it was smudged or printed poorly and wouldn't be recognized in normal circumstances.

In this way neocortex permanently produces the future on every level of its hierarchy. On highest conceptual levels we also busy with predictions: "Who is going to walk through this door? "What someone is likely to say next? "What I expect to see when I turn the corner? "What wold be the likely results of my own actions? and so on.

At that we are making mistakes quite often, mostly because our brain by its construction identifies absence of disproof or nonconfirmation with confirmation[26], hence our recognition thresholds are lowered by default. Thought,

[24]We will come back to this important feature of human brain later, when we will talk about invariants and universals

[25]Here Kurzweil alludes to the central idea articulated by Charles Lyell in his main work on geology "Principles of Geology"(1830) from which Darwin heavily borrowed when elaborated his own theory.
Kurzweil, R. How to Create a Mind. Penguin Publishing Group. London. P.116

[26]You do remember priority of speed over precision, don't you? We must also keep in mind that absence of information "costs"much less than information about an absence.

by and large, could be reduced to sending somehow consciously manageable inhibitory signals down to our lower level recognition modules thereby forcing them to ask for more data to confirm from their "subordinates".

Andrey Kurpatov makes the same point: "So, for example, if I consider some state a "potential adversary"I necessarily interpret its political actions as hostile, and i will need to make a huge, focused effort to discern in these actions something different."

Neurophysiology also agrees that perception can not be explained as strictly upward process. We rather have some sort of recursive "hermeneutic circle"of hypotethis generation on upper levels and their testing on lower levels. According to this kind of theories of perception percieving subject is building her actual, constantly renewing model of the world in cycles. In such a cycle one's current needs, expectations and interests shape hypotheses for one's perceptual systems to confirm or disconfirm, and a rapid sequence of such hypothesis generations and confirmations produces theu ltimate product, the ongoing,updated "model"of the world of the perceiver.

We would like to emphasize once again that *needs* must be understood in maximally abstract, *phenomenologica*l sense - as intentionality, some specifically shaped interest, "being-directed-toward..". A. Kurpatov also draws our attention to this: "It would be weird to speak about "need"in mathematical calculations, or about "need"in search for truth, or about "need"in music. But nevertheless, within the framework of internal psychic mechanics itself this is actually the "needs": solving of intellectual problems - mathematical, philosophical, psychological, problems of social communication or of aesthetic pleasure and so on - for the mind they all do not differ from solving problem of satisfying physiological needs.[27]

Prominent and highly honored by me researcher of consciousness Daniel Dennett points out that mentioned earlier generate-and-test theory of perception

[27] *Kurpatov A.* "Thought. Systematic investigation". - SPb, 2019. P. 55-56

explains very well not only normal brain functioning, but also sheds some light onto the nature of hallucinations: all we need here for an otherwise normal perceptive machine to be thrown into hallucinatory mode is for hypothesis-generation (expectation-driven) side to operate normally, while perception (data-driven) side of the cycle goes into random or arbitrary round of confirmations and disconfirmations. In other words, if noise in the perceptive channelwould be arbitrarily amplified into"confirmations "and "disconfirmations"about intellectual object on cernain level being percirved, the current expectations, concerns, obsessions, and worries of the subject will lead to framing questions[28] or hypotheses whose content is guaranteed to reflect those dominant interests, and so a quite freakish perceptual picture might unfold in the perceptual system, strictly speaking, without an author.

To elaborate on this theme we could hypothesize that notoriously famous Elizabeth Loftus's "false memories"might have the analogous structure - as soon as our main concern is not the past, but mostly the competitive future it is quite natural to assume that we must be interested in images from the past merely as in "lessons of history"of some sort, pieces of which might be useful in building this future. It's no wonder then that some parts of this construction would be assembled, as one might say, "quick and dirty": "It would be unreasonable simplification to call this process "mirroring"as soon it is not mirror image of what's real, but sort of it's recreation, production."[29]

Let's turn back to introduced earlier function of expectation. French philosopher Alain Badiou uses the same mathematical construction when he tries to to make a distinction between ontology and phenomenology and speaks about *logic of appearance*. He illustrates very gracefully his main thesis with the example of Paul Dukas's opera "Ariadne and Bluebeard". First of all, Badiou notes that precisely proper names "Ariadne"and "Bluebeard"make

[28] `https://en.wikipedia.org/wiki/Framing_effect_(psychology)`
[29] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 72

possible appearing by means of discrete, generally speaking, situations (musical, scenic or narrative): Ariadne before knowing Bluebeard, the first encounter between Ariadne and Bluebeard, Bluebeard the murderer, Bluebeard the child, Ariadne freeing the other wives, Bluebeard and Ariadne' sexual encounter etc. Appearing of characters of these two is in no way regulated by, as Badiou puts it, "genealogical constructions required in order to fix within the real the referent of these proper names". Just as the peripeteia that affect Bluebeard and Ariadne presuppose that material referents of their names remains the same, but this "same"does not ever appear, being strictly reduced to the names.[30] Appearing, on the other hand, grasps only in a form of series of transcendental values which measure the extent of self-difference, as well as the difference from the others in different situations (love, sex, death, the dream of freedom) and amounts to explicating in which sense, situation by situation, Ariadne is something other than 'Ariadne' as captive of Bluebeard, Bluebeard something other than 'Bluebeard' than a sinister maniac, but also how Ariadne is something other than Bluebeard's other wives, even though she is also one of them. It's quite obvious that when we evaluate the last difference we must take into consideration the fact that Ariadne, unlike other wives, is clearly aware of her position and does not want to accept it. Other five of his wives, on contrary, can be substituted with one another in their relationship to Bluebeard and in this sense they are "transcendentally identical which is what marks their 'choral' treatment in the opera. Following the same logicr of appearing, we immediately know how to evaluate Bluebeard in love with Ariadne and in result finding it impossible

[30] *Badiou A.* Logics of worlds. N. Y. 2009. C.116. It is interesting to compare this thought with the Kurpatov's text selection: "If they tell me: "Saint Jerome"or "Great Caravaggio but there is no neither these intellectual objects in my individual world of intellectual function, nor adjacent to them - such as "religion "saints "artists"with their possible "greatness appropriate information simply won't be accepted, it would be like talking to the brick wall. *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 94-95

to treat Ariadne like the other women, and thus stands outside of what is implied by the referential being of the name 'Bluebeard'. Bluebeard is not "Bluebeard so value of function Ess(Bluebeard) $= \mu$ i. e. minimal. Again, within the opera this limit value corresponds precisely to the extravagant fact that for the duration of the last act Bluebeard remains on the stage, but does not sing or speak a single word: Bluebeard is absent from himself.

All of the above gives hope that given mathematical model which has been chosen for the concept of intellectual object to be formalized has been chosen correctly and quite fully satisfies the requirements for it.

So,from now on, by intellectual object we will understand quite intricately constructed, synthetic object $\mathbf{A} := (A, \mathrm{Exp}_A)$, comprising set of data A and function of expectation $\mathrm{Exp}_A : A \times A \to \Omega$, which depends essentially on the subject of experience and her internal state. In mathematics such object $\mathbf{A}$ called *Heyting-valued set*, or, since values of function Exp_A belongs to Ω - an Ω-*set*, while the collection of all objects of this type endowed with some additional structure transforms this collection into the *category* of Ω-*sets*. Study of this additional structure is what we will do next.

II. Intellectual function

"By itself, every intellectual object
is of course a set...
The relations between these sets
this is what we call "intellectual
essentially, mathematical function.

Kurpatov A. Methodology of thought.

We saw that intellectual objects are not some chaotic, messy collection, but they form complex hierarchy, i. e. it is a highly structured collection with many varied relationships found between it's elements. Following the author of the quotation given in the epigraph to this part of our study we will call relation between intellectual objects *intellectual function*. According to Kurpatov, the action of this intellectual function is the only instrument of thought: "We operate endlessly with these operands (intellectual objects) inside our own head producing in that way more and more relations between them. These new relations are in fact new - *derived from* - intellectual objects."[31] And further: "When we say we are "looking for understanding"of something, in reality we communicate our desire to create an intellectual object which will be a solution to the problem we are occupied with at the moment."[32]

Thus, our immediate goal is to outline some generalized space for the consistent discourse, where we could control actions of intellectual function. As a model of such generalized space we will examine a category - algebraic structure loosely can be described as a collection of "objects"that are linked by "arrows". [33]

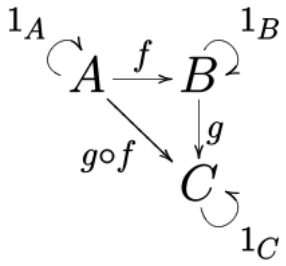
[31] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 37
[32] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 117
[33] `https://en.wikipedia.org/wiki/Category_(mathematics)`

If we will think of the relation as a some oriented connection from one object toward another, it seems quite natural then to assign letters A, B, C... to intellectual objects, and arrows - to relations. It doesn't mean though that there can not be bilateral relations - but in such cases we will think that there are two oriented connections between these two objects. And in order for the rather vague concept of "relation"to give a more rigorous look - of, albeit intellectual, but still a *function*, we will require that we could formally manipulate with the arrows (relations)more or less like with the ordinary mathematical functions: for instance, it would be convenient if we could concatenate them together just as we do when adding numbers. Just as the result of adding two numbers is a number, i. e. an element of the same set where operation of addition has been determined, we might want the result of concatenation of two arrows f and g (which is, by the way, customary to call their *composition* and to denote $g \circ f$) to be an arrow that also belongs to given category.

To continue the analogy with the addition we will also require an analogue of zero - so called *identity arrow* (id_A, or 1_A), concatenation of which to any other arrow (relation) does not change the result of the relation. There is a slight difference though: since the arrow is determined not only by it's "internal nature but it is also important to specify where from and where to given arrow is going, there has to be as many identity arrows as there are objects in the category. And one more thing: as soon as in this case we add not numbers, but "(inter)actions the order of actions is also very important: hence, generally speaking, $g \circ f \neq f \circ g$. Respectively, when adding identity arrow we must keep in control the side, which identity arrow is added to. In other words, the fact that concatenation of identity arrow does not change the result of the relation must be formalized as follows: for every relation $A \xrightarrow{f} B$: $f \circ 1_A = 1_B \circ f = f$.

Here is an example of one of the simplest formal constructions which

consists of only three objects, but which is still a category:



Notwithstanding identity arrow might seem "superfluous it serves a number of important roles. For instance, it would be quite difficult to state formally many basic categorial definitions - such as *isomorphism* of objects[34], to name one: two objects A and B in category C called isomorphic if there are two arrows $f : A \to B$ and $g : B \to A$ such as $g \circ f = 1_A$ и $f \circ g = 1_B$.

Now let's recall that the objects **A**, **B**, **C**... which we want to "connect"with the arrows are Heyting-valued sets and thereby they have quite rich structure - actually, they themselves are the result of certain relation - relation of the form $f : A \times A \to \Omega$, to be precise. So, basically we must define functional relationship $r : \mathbf{A} \to \mathbf{B}$,that in itself is the relation between relations, which means that this type of connection must be additionally restricted, so that these restrictions would take into account the internal structure of both set **A** and set **B**.

If we will think of function $\mathrm{Exp}_A : A \times A \to \Omega$ as of encoding of information about subject's system of differentiations which she discovers in the situation, it becomes quite obvious that relation $r : \mathbf{A} \to \mathbf{B}$ also must depend on this subject of experience Ω and must take into consideration somehow this information. If r would be an arbitrary function it would simply

[34]In simple words, isomorphism is identity of structures. In not that simple words, isomorphism is invertible morphism. (Morphism is another name for the arrow in category.) That is, isomorphism is a type of relationship in which only objects very similar in something can be.

be a rule establishing correspondence between elements of sets A and B and we could fix the fact of such correspondence as equality $r(a) = b$, where $a \in A$ and $b \in B$. Now, to stay strictly we could interpret this equality in the same terms of expectation, i. e. to assign more or less high values $q \in \Omega$ to all couples $(r(a), b)$, depending on how $r(a)$ close to b. Detailed formal definition requires quite a lot of technical work[35], but we will skip it here, since category of complete Ω- sets allows for mush less technical and much more meaningful interpretation.[36] As we already mentioned relation between intellectual objects must take into account the information contained in these objects which can be basically reduced to certain system of differences and identities. In other words, relation *must respect* the distribution of differences and identities that has been already done by function of expectation Exp, i. e. $r : \mathbf{A} \to \mathbf{B}$ can be represented as ordinary set theoretic function which conserves all those differentiations and identities. It does not create anything - neither increase identity of element with itself, nor increase it's difference from the others:

$$\forall a, b \in A : \qquad \mathrm{Ess}_B r(a) \leq \mathrm{Ess}_A a$$

$$\mathrm{Exp}_A(a, b) \leq \mathrm{Exp}_B(r(a), r(b))$$

At last, we have a minimal set of instruments which we may use for modeling of action of intellectual function as a process of building of more and more complex intellectual objects. Andrey Kurpatov suggests to break down this action into two large stages:

1. "On every level of mentality - from the elementary perception to the most complex intellectual reasoning - one and the same intellectual function

[35]Look for details in Goldblatt R. (2006) Topoi. *The categorial analysis of logic.* P. 277-278

[36]Borceux, F. (1994) *Handbook of Categorical Algebra*, volume 3, P.160. Why exactly space of thought can and must be identified with the category of complete Ω- sets we will try to substantiate below.

of the forming (reproducing) of intellectual objects is performed. We identify them as objects (from the variety of stimuli something is emerged - some "thing") and then we correspond them with the rest of mental content (other intellectual objects) which is already there."[37]

2. "Perceived "thing"(unit of thought) being subject to complexity - as if it is raised to the power of those knowledge (intellectual objects) that we have, and gets a corresponding value for us - considerable, in case there already are many intellectual objects that being involved by our intellectual function into this process of producing a new intellectual objects, and marginal, in case there are no such intellectual objects in us, or they rendered irrelevant by our intellectual function. [38]

In the language of category theory these stages correspond to the following categorial operations:

- Finding the categorial limit A of the piece of category that contains the set of *relevant* intellectual objects, i. e. those which were involved by intellectual function into the process of producing a new intellectual object B ("perceived "thing").

- Exponentiation, or raising of perceived "thing"A to the power of B, i. e. construction of so called exponential object B^A, which, in turn, also can be considered as categorial quasi-limit. [39]

Important note. Usually, when we perform basic arithmetic operations - for example, with integers - we are not bothered with the question why the result of this or that operation exists at all, or "where"it exists. We take it for granted that the result of addition, multiplication, subtraction or division of numbers is also a number. Although even after subtracting a larger number from a smaller number, people were not immediately able to ascribe any meaning to the result they obtained. Strictly speaking, in order to find the

[37] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 95-96
[38] Ibid. P. 96
[39] *Badiou A.* Mathematics of the Transcendental. N.Y. 2014, P. 49

result of this operation within the same set of elements, it was necessary to "throw in"negative numbers, as well as zero, into the known and familiar set of natural numbers - in other words, in order not to leave the set, on which the operation of subtraction is defined, and elements of which we call numbers. If, in principle, some operation defined on the elements of a set does not take us outside this set, then such a set is called *closed under* this operation. When we introduced the operation of composition of arrows of the category we encountered the same problem - we required that the result of "addition"of two arrows in category would also belong to the category. Right now, when we talk about taking the limit or constructing the exponential object, we also need to make sure that these objects exist in our category. The category that is closed under those two operations called *cartesian closed*.[40]. It has been already mentioned earlier in another our work[41] that in this way it is possible to point out implicitly on the size of the category: if we use optical metaphor then the category would be the bigger the more objects it has, from which we can *see* different pieces of the category. In this sense given category of Ω-sets is sufficiently large.

Categorial construction that includes the object C from which all other objects of a given piece of category are "visible"called a *cone*. Piece of a category in this case called a diagram. "Visibility formally speaking, means existence of arrows which go from object C towards all objects of the diagram so that every arrow of the diagram enters into the composition of an arrow coming from C. In result the whole construction will *commute*, i. e. if there is more then one way to reach one object from another, these ways are equivalent, or interchangeable.

Given diagram may admit several different cones, so we will say that the diagram admits the *limit* if there is the object C in the category, from which

[40]Badiou, A. *Mathematics of the Transcendental.* N.Y. 2014, P. 51

[41]*Egorychev I.* Categorical analysis of Kurpatov's "Tractatus psychosophicus"in context of AGI perspective developement"

we see the diagram as closely as possible, i. e. if there is another object C_1 from which the diagram is visible, then C is also visible from it.

It is quite evident that one can't get all necessary relations (arrows) to form the cone unless the limit object would have specific structure: indeed, such object would contain all the information about given piece of category - no more and no less. and in this particular sense the limit object is *optimal*. Thereby finding of the limit object A for some diagram inside a category is equivalent to construction of the heavy intellectual object, according to Kurpatov - extension of our knowledge, relevant to the problem we are working on at the moment, forming it into the "some thing"to the power of which a newly discerned "thing"could be raised: "In fact, every single action of intellectual function is further sophistication of intellectual objects. That is in reality I does not just find some brand new relations between those intellectual objects I already have, but rather form new intellectual objects transforming into them a number of objects in question."[42]

It is very important to point here at the difference between diagram as an arbitrary piece of category and subcategory - the latter, together with relevant (intellectual) objects, comprise all! relations that exist between them in the category. [43] And here lies a fundamental difference between presentation and reconstruction. Kurpatov mentions on the subject: "When, through the rupture of our representations, we see maximum possible number of facts we are not perceiving anything particular, but rather something given as a relation, and not as a relation between this and that, but as a result of relation - essence of the relation. [44]

That is, this probable absence of some, possibly, very important relations (arrows) in the diagram is precisely what allows us to identify it both with our representation in general, and with those *essences*, the representation consists

[42] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 116-117
[43] `https://en.wikipedia.org/wiki/Subcategory`
[44] Ibid. P. 65

of. "Bringing the "essences"into its "objects"is normal and necessary practice for our mind". "Relation is at the core: following our needs we enter into a relationship with actual reality trying to satisfy somehow these needs. And just as a result of these relations we form corresponding "essences"which, ideally, will stay invariant for many different phenomena potentially capable to satisfy one or another our need (want)."[45] Pointing out on this deficit of relations in the diagram that prevents to turn this piece of category into subcategory, we thereby want to emphasize the biased nature of every mental representation, its, so to speak, "utilitarian"character and, as a consequence - inevitable distortion of reality by its constituent idiosyncratic entities's "gravity field".

So, we are dealing here with two radically different types of action of intellectual function:

- exponentiation, which corresponds to the sophistication of intellectual object B by means of raising it to the power of existent knowledge/representations A;

- constructing of so called hom-functor $\mathrm{Hom}_C(-, B)$ which puts every object A that belongs to subcategory $C \subseteq \mathbf{Set}$ (composed of relevant objects of given diagram and supplemented with all missing relationships between them) into correspondence with the *set of all relations*[46] between object B and object A.

Once again, we would like to draw attention to how different these two actions are: while the exponential object is the object that lies in the initial cartesian-closed catregory $\Omega-\mathbf{Set}$, hom-functor in fact embeds our subcategory $C \subseteq \Omega - \mathbf{Set}$ into another category of all sets **Set** that is significantly richer both with objects and relations - which gives us opportunity to investigate

[45] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 76

[46] In many categories arrows represent special kind of functions called homomorphisms - a structure preserving maps. Hence, the name of the corresponding, specifically constructed hom-functor

given chunk of knowledge not in the isolation, but as a part of a much larger network of relations, which usually can not be discerned from within the initial category.

Moreover, the reconstruction of much higher level is possible - the one that puts every intellectual object A of subcategory C in correspondence with functor $\mathrm{Hom}_C(-, A)$ which contains allthe information about relations of object A with every other object that lies in C. If this is the case, then subcategory C will be embedded in even more saturated network of relations resulted also from exponentiation of some sort - only in this particular case category of all sets **Set** must be raised into the power of our subcategory C! As a result we will get so called functor category, objects of which are all functors from C to **Set** and arrows between them are *natural transformations* of functors.[47]. Possibility of this construction is guaranteed by remarkable mathematical result known as *Yoneda embedding* [48].

[47] *Egorychev I.* Language of category theory and "the limits of the world". SPb. 2018, P. 57-58

[48] Ibid. P. 109-112

III. Thought: invariants and univelsalia

"Invariants are the "words"of
non-referential language."
Kurpatov A. (From personal correspondence.)

So, we have just seen that mere sophistication of intellectual object rather looks like representation then reconstruction and thereby does not guarantee us the elimination from fragmentary and biased nature of the former: "Factual reality can't be given to us in representation, because represen tation is always distorted by the agent who creates this representation... In fact, all our representations of reality are value judgments, while the reality itself could be only theoretically reconstructed."[49]

But how exactly in this case elimination from the aberrations of "reality effect"created by every single representation of it, is possible in practice? In other words, how reconstruction is possible? The answer, as it seems to us, lies in following, remarkably deep idea, also delivered at due time by Andrey Kurpatov: since intellectual objects are *virtual machines*, any hierarchy of them has virtual, or we would rather say, *conceptual* character as well. This purely abstract level, on which some particular intellectual object is seated, even can be calculated at your wish by counting a number of those "Aha!-effects"which were necessary for this intellectual object to form.

"We need to stop thinking of "levels"of thought and thereby to get rid of delusions that arise because of this - Kurpatov insists. Differences between various kinds (types) of intellectual objects that we discover mat seem enormously huge: we have "feelings"here and "abstract numbers"there,

[49] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 53

they say, "physical regularities"here and "metaphysics of existence"there. But all these are utterly insignificant for the thought as such. [50]

Consequently, thought, supposedly constrained by specific content of at times radically different contexts (remember once again the "slow gentle force"of Lyell and Darwin), inside its own activity proves to be independent of them. As Kurpatov wittily notes: "There is no problem for thought to measure "feelings"with "abstract numbers or to find "metaphysical specificity" in "physical regularities". [51]

The freedom of thought thus obtained brings us to the most important concept of the methodology: that of *invariant* - extremely important instrument of thought which allows us to think disregarding specificity of particular contents.

As soon as thought is the process of construction of more complex intellectual objects from more simple ones carried out by intellectual function, then invariant also has to be an intellectual object of some sort. But what sort? We already mentioned above that virtual machine of mind can't work in isolation - it needs certain ecosystem where other auxiliary virtual machines can be uploaded. In this sense human brain essentially depends on society and on presence of language in it, which is the main instrument of building of such ecosystem: vast amount of knowledge coded in a form of texts, charts, schemes etc. uploads into the environment. However, it is by no means necessary for the virtual machine of general intelligence to be linguistic machine - moreover, in case of artificial intelligence such a restraint would seem probably too "artificial". But equally probable that it would be almost impossible to implement brain principles programming without one or the other quite elaborated system of symbols.

It means that we need to supply our cartesian closed category (which

[50] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 92
[51] Ibid. P. 92-93

from now on we will consider a "world of representations or "my scheme of the world"in terms of "Tractatus") with some quasi-linguistic structure - a special system of signs, or names which can then be effectively (and what is most important - disregarding specificity of particular contents) used by any user who would thereby be able to deconstruct his or her representations and to theoretically reconstruct factual reality with the help of suitable invariants created individually for the specific tasks.

At the core of this process lies our ability to classify things. As A. Kurpatov wrote: "A "table"serves us as invariant as well as a "face"(or a "cup"or any thing of that sort): anything we can assign the function of the table to can be identified as a "table while any thing we can assign the function of the face to can be identified as a "face although this thing in fact may not be the face". [52]

Actually, pretty common procedure has been described here which we carry out constantly without even thinking what we doing: for instance, I can't add "tables"to "chairs"or "pears"to "apples"unless I bring them to the common denominator "furniture"or "fruits". "Furniture"and "fruits"in their turn are also don't "marry"each other so well unless they will be grasped by common concept of "objects"while "objects"and, let's say, "feelings"can't be perceived as something even remotely "similar"before I start to think of them as of "things". The certainty that is lost in the process of such an ascent and the freedom thus gained at the same time, are both perfectly captured in English with the help of *indefinite* article: a cat is not *this* or *that* cat, it is *any* cat whatsoever.

But it's still not enough - in order to disregard referentiality we need such a system of symbols where there is no a single signifier would correspond to any thing in my scheme of the world. That is why from now on we will proceed from important methodological assumption: invariant is a *floating*

[52] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 75

signifier.[53].

It is very important to note here that different "users"(subjects of knowledge) would necessarily have different signifiers which they chose to play role of invariants, because the latter essentially depend on modus of existence of particular user (class of users) - they will be "colored"by this modus of existence, so to speak.

Invariants are what allows the knowers to reconstruct something regardless of who exactly uses them: human, machine, alien, amoeba. Universalia, on the other hand, are what gives reality its certain positive characteristics depending on how this particular knower is organized. In other words, while invariants are invariant for all beings, universalia are universal (invariant) for this or that particular being (knower).

This is what A. Kurpatov writes on the subject: "We think by means of these universalia - big and small, heavy and distant, thin and complex, empty and huge, bright and strong, singular and multiple, heap and couple, tall and powerful, petty and insignificant, abundant and overfull, fast and slow, continuous and instant, space and territory, zone and region [S. Pinker]. In other words, it is our modus of organizing things: our objects, and all of them are intellectual, are given to us that way - they are the way they are by default. [54]

Notoriously famous concepts of Yin and Yang, I think, colud provide good examplificaion of invariants: maximally derived from referentiality these principles exhibit themselves as soft and hard, male and female, light and dark, weak and strong, submission and domination etc. Interestingly enough that concepts of "strength "female "light"or "submission"themselves can be considered as invariants of certain level, depending on purposes they are used. In this "metasense"concept of "universalia"also could be thought as one of

[53] `https://en.wikipedia.org/wiki/Floating_signifier`
[54] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 110-111

the invariants.

That is, we always will understand invariant in two senses:

- in narrow sense - as in cases of the concepts "a cat"or "Yin when individual instantiations do not change the concept;

- in a broad sense - as an methodological instrument of reconstruction of reality, independent of particular knower by definition.

It is clear that in the latter case we can not know in advance whether this or that floating signifier will prove to be invariant or not. We only can verify it "empirically so to speak: if both I and, let's say, a computer will get *equivalent models* of reality, then both I and "her"used invariants, which will work as such in this case strictly by definition.

To sum up all've been just said, in order to model the world of intellectual function correctly, we will need to supplement our structure with the flexible system of signifiers using which any user could change more or less freely its hierarchical order and to define them through each other depending on the problem that is being solving at the moment. What kind of signifier will be chosen specifically as invariant will also depend on the domain where the problem was posited: "Every area of expertise is determined... by its own specific "entities"which, in their turn, have their own logic of relations and specific limitations appropriate to this area. That's why every such "private"area would require specific invariants, which, most likely, couldn't be expressed otherwise then in specific, complementary to this area universaia.l[55]

There is one more quite important moment: just mentioned flexibility of the system of signifiers brings with it a number of potential costs. In particular, a freedom gained through such flexibility resurrects so called "the theoretician's dilemma"well known in methodology of science - thesis that problematize both the role of the theoretical terms in scientific theory and of concepts with so called "unobservable entities"as their referents. In

[55] *Kurpatov A.* "Methodology of Thought. A Draft". SPb, 2018. P. 172

other words, in extreme we obtain an instrument symbolic resources of which threaten to significantly exceed existential resources of being. After all, neither the famous dialectic principle of the unity and conflict of opposites nor any other arbitrarily pretentious axiomatics founded by such invariant beliefs as, let's say, "All is ashes"or "Tis love, that makes the world go round"are in no way inferior to our example of Yin and Yang.

Kurpatov repeatedly emphasizes that in the essence of any area of expertise or practice we must discover something *real* - something that is really exists. He is deeply convicted that we can't build a working methodology based on such concepts as "god "ethics"or "history because with all desire we won't be able to answer the question: "What is it *really*?"In this sense invariants "intellectual object"and "intellectual function"are rather observable because at least in principle could be reduced to the "real machine"of neural assembles, logic of their connections and transmission of nerve impulses.

Thus, strictly methodologically speaking, invariants can and must describe *what* really is, while universalia describe *how* it is to someone (in relation with someone).

It was exactly the latter important moment that was successfully formalized in such algebraic construction as Heyting-valued set. In it, we recall, the subject of experience is "grasped"in two aspects - in respect to primordial differentiating ability, i. e. only as a Heyting lattice Ω, and as a result of relation, i. e. as a thing ("the world that presented itself to me as me in this particular way")$\mathbf{\Omega} = (\Omega, \mathrm{Exp}_\Omega : \Omega \times \Omega \to \Omega)$.[56]. We wrote above that conditions imposed on function of expectation allow for considerable variability of the latter, however this variability vanishes when we construct

[56]For details see: *Egorychev I.* Categorical analysis of Kurpatov's "Tractatus psychosophicus"in context of AGI perspective developement"

Ω as Heyting-valued set - this function must be defined in a unique way

$$\mathrm{Exp}_{\Omega}(p, q) =_{def} p \Leftrightarrow q \quad (1)$$

Operation of equivalence $\Leftrightarrow$ that has been just introduced is derived operation on Heyting lattice, and corresponds corresponds, as expected of equivalence, to two implications,i.e. $p \Leftrightarrow q =_{def} (p \Rightarrow q) \wedge (q \Rightarrow p)$, where implication $p \Rightarrow q$ in its turn is short notation for so called relative pseudo complement $p \Rightarrow q =_{def} \bigcup_{x \in \Omega} \{x : p \wedge x \leq q\}$. We remind that when "absolute" pseudocomplement is a topological analogue and convenient generalization of negation in logic, relative pseudocomplement is a generalization of logical connective "if, then..."(To advanced readers we suggest to check themselves that absolute pseudocomplement really is a pseudocomplement relative to $\varnothing$, or minimal element of Heyting lattice.)

It means, as we already mentioned before, that having at our disposal a srtucture with partial order and a number of "order"analogues of set theoretic operations of intersection and union ($\wedge$ и $\vee$ - greatest lower and least upper bound, or meet and join)[57] makes possible for us to mirror almost completely the structures of rationality as such.

So, informally, formula (1) only says that elements are close to each other to the extent that they are equivalent. Or, to put it differently, two elements are maximally close, only if they are equal. Which might probably look tautological, in a sense, but it can't be otherwise. Indeed, degree of proximity of two elements that belong to "me as a thing"which simultaneously serve as the marks of my differentiating scale, i. e. values of the scale itself must be such that the structure of me as a subject of experience would remain intact.

[57] Randomness of the structures that only partially ordered can be increased indefinitely. Condition for any two elements having greatest lower and least upper bounds guarantees that randomness of our existential ruler won't be too large. For details see: *Egorychev I.* Language of category theory and "the limits of the world". SPb. 2018, P. 28

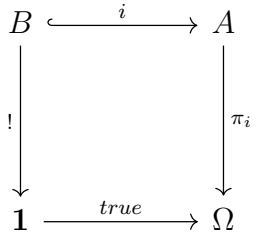
From such rigid and quite tricky from technical point of view specification of the function Exp_Ω, however, follows everything we need[58]: and in particular, that the proximity of an element p to its own "essence"is always maximal, that is $\mathrm{Ess}_\Omega(p) = \top$. Curiously enough, given property can be interpreted so that I cannot doubt that something seems to me different and how different this something seems to me - just as Descartes once discovered the doubtlessness of his own doubt, so do we, following him, may postulate: "I discriminate, therefore I am."

In "Categorial analysis of "Tractatus"and in more detail in "Commentaries to "Analysis"we have already expressed the general idea of how a category, which at the same time is a topos, manifests a much richer structure, which has a distinctively logical (symbolic) character. Moreover, this manifested "internal logic of topos"generally is not Boolean one, and depends, generally speaking, on which particular topos we are considering (in fact, that is why this logic is called internal). Despite the fact that above is true for every topos as Cartesian-closed category with subobject classifier, a world of intellectual function considered as category of Heyting-valued sets will have a number of important distinctive features, which we will now study in more detail.

So called Ω-axiom is formulated in exact correspondence with our intuitive view that in order to single out in a given object A a certain part i of it we need *to point* on some attribute π_i which we will use for specification of that part - and for that, in turn, we will need signifiers (signs, symbols, names):

for any subobject $i : B \to A$ there exist unique arrow $\pi_i : A \to \Omega$ such that following diagram commutes:

[58] *Goldblatt R.* Topoi. The categorial analysis of logic. N. Y. 2006, P. 277



If in a category C there exists an object Ω satisfying given axiom, then together with an arrow $true : \mathbf{1} \to \Omega$ this C-object is called a *subobject classifier*, or a *truth value object*. [59] A truly remarkable fact is that in the category of Heyting-valued sets such object really exists and, moreover, it turns out to be a Heyting valued set $\mathbf{\Omega} = (\Omega, \mathrm{Exp}_\Omega : \Omega \times \Omega \to \Omega)$![60].

It follows from what has been said, first, that the world of intellectual function is really a topos. And second, the role of subobject classifier that is the role of an instrument which is used, from the one hand, to single out in any intellectual object any part of it and, from the other hand, to name every singe part, is played by Ω itself. It means that not only the world of intellectual function is structured by my (non-convertible) values (meanings), but I act in it as the "measure of all things or truth value object. Formally it can be put as follows:

$$\forall A \;\; \exists \theta_A : \mathrm{Sub}_C(A) \cong \mathrm{Hom}_C(A, \Omega),$$

i. e. for every object A is true that to each of its subobjects some name corresponds, and different subobjects correspond to different names (symbols). In other words, for every intellectual object as well as for every arbitrarily small part of it we get sought ramified and flexible system of signifiers, which is much more convenient to operate with. And as soon as there is a natural

[59] Object of category C is called *terminal* if from any other object of category C toward this object there is only one arrow. *Subobject* is a categorial analogue and generalization of set-theoretic notion of a part, or subset.

[60] Goldblatt R. *Topoi. The categorial analysis of logic.* N.Y., 2006, P. 277

isomorphism between the parts of objects and the names (propositions) that are used for specification of those parts (remember that every object at the same time is its own part), both when taking limits and when exponentiating, a completely controlled conversion from objects to their signifiers and back is possible.

And, finally, third: let's try to understand how the procedure of signification is carried out in the category of Geyting-valued sets. In "Categorial analysis of "Tractatus"we said that by predicate $\pi(x) : B \to \Omega$ it is most natural to understand a function which associates every element x of set B with some degree of its proximity to the attribute expressed by given predicate, or a certain truth value on the scale Ω - a degree to which the proposition "element x has property π"is true (according to our current interpretation - the subjective measure of our expectation of how fully a given element x will exhibit a given property π") Hence, sometimes it is customary to denote maximal element of the scale as $\top$ (true), and minimal - as $\bot$ (false). Let's not forget that we gain much more freedom than in two-valued logic due to the presence of intermediate, and often incomparable truth values in Heyting algebra Ω.

Since $\mathbf{B}$ is not *just* a set, but an object of our category ($\mathbf{B} = (B, \mathrm{Exp}_B)$), it is naturally to demand from predicate function to be consistent with those expectation values that depend on the function Exp_B. In particular:

- it is impossible for an element x to exhibit a certain quality (property) and at the same time to "look like"another element y in a degree higher than a degree of exhibition of that quality by the element y: $\pi(x) \wedge \mathrm{Exp}_B(x, y) \leq \pi(y)$;

- it is impossible for an element x to exhibit a particular quality (property) in a degree higher than a degree of manifestation of it's own essence: $\pi(x) \leq \mathrm{Ess}_B(x)$.

If we add to these two conditions the third one:

- $\pi(x) \wedge \pi(y) \leq \mathrm{Exp}_B(x, y)$ (which means that a degree in which elements x and y at the same time exhibit quality π, can not be higher than a degree of proximity of those elements), then we will see that given predicate π singles out in object B the part with *no more than one* element exhibiting the property π with maximal degree. Indeed, if $\pi(x) \wedge \pi(y) = \top$, it follows that $\mathrm{Exp}_B(x, y) = \top$, i. e. such x and y are identical!

Such "atomic"part is a Heyting-valued analogue of subset consisting of one element x and called *singleton*. Accordingly, it is natural to call a respective signifier, or predicate function π which satisfies to these all three conditions, also a singleton, or *atomic signifier*.

Now recall that the function of expectation $\mathrm{Exp}_A : A \times A \to \Omega$ assigns to every pair $x, y \in A$ a certain degree of their coherence q on Heyting algebra Ω. If we'll keep one element of the pair fixed (let it be $x = a$, for instance), this function wil turn into ostensive definition - it will be function of one variable that assigns to every element $x \in A$ a degree of it's proximity to fixed element a. Such function is a predicate - in this case we express the property π by pointing at the particular representative a, or at some archetypal exhibitor, so to speak, of the property in question as if to say thereby: "it is like that a- we will call it *ostensive* predicate and denote as $a(x)$. Moreover, this function will be atomic by obvious reasons.

Correspondingly, Heyting-valued set **A** is called complete Ω-set, if all its singletons are of the form $a(x)$ for some *unique* $a \in A$. [61]

And now attention, the most important: remember - we said above that by invariants must be described somthing that really exists. But what is this if not the requirement of completeness of Ω-set A: at least atomic signifiers must point at some thing that exists in reality. So we argue that taking into account what has already been said above it would be the most correct and the most appropriate to describe the world of intellectual function as

[61] Goldblatt R. *Topoi. The categorial analysis of logic.* N.Y., 2006, P. 388

a subcategory in category of complete Ω-sets, which from now on we will denote as **CΩ-Set**.

As we already mentioned, in category **CΩ-Set** many objects and relations could be defined in more graphic and direct manner. That was particularly the case with the definition of arrows: instead of quite awkward and technical construction in the world of intellectual function we agreed to call a *relation* a function $r : A \to B$ defined on usual sets, such that:

$$\forall a, b \in A : \qquad \mathrm{Ess}_B(r(a)) \leq \mathrm{Ess}_A(a)$$

$$\mathrm{Exp}_A(a, b) \leq \mathrm{Exp}_B(r(a), r(b))$$

From these two conditions immediately follows that $\mathrm{Ess}_A(a) = \mathrm{Ess}_B(r(a))$. That is, when engaged in a relation an element a of the set A remains close to its own essence in the same degree that preceded such engagement. Conditions of this kind are called conservative which means that relation as a function *conserves* (does not change, leaves intact) initial distribution of values on Heyting lattice that have been made by the function of expectation - it *respects* them.

Now let's examine in more detail some other important constructions in **CΩ-Set**.

We'll start with terminal object. Earlier we defined a terminal object as marked object of category C such that there is only one arrow into this object from any other object in C.[62] It is denoted as **1**, because in the category of sets it corresponds to singleton - one element set $\{x\}$. Examination of constitution of terminal object in category **CΩ-Set** is of interest, also because this relatively simple example is a very graphic illustration of how greatly

[62] Almost every important categorial construction allows for its dual - which may be obtained from the initial construction n by reversing all its arrows. Such dual to terminal object is so called *initial object* which has exactly one arrow from it into any another object of the category.

formally identical categorial constructions can differ in different categories.

The specific of Ω-set's construction plus conditions imposed on arrows in the world of intellectual function imply those specific restrictions that organize the shape of terminal object in it: so it's almost no surprise that the role of terminal object **1** in this case will be played again (as in the case of the truth value object) by subject of experience Ω together with the function of expectation defined as $\mathrm{Exp}_{\mathbf{1}}(p, q) = p \wedge q$. Due to existent restrictions on arrows there willbe exactly one arrow into this complete Ω-set from any other Ω-set **A** that assigns to each $a \in A$ the degree of its essence $\mathrm{Ess}(a) \in \Omega$.

By the way, an element $a \in A$ such that $\mathrm{Ess}(a) = \top$, we will agree to call a *global* element. In this context it is also make sense to consider another function which sends all elements $a \in A$ into the minimal element μ in Ω. [63] This function will be a singleton (see the definition above), but it means that in intellectual object **A**, which is a complete Heyting-valued set, as we remember, there must be a unique element $\varnothing_A \in A$ such that $\forall x \in A \quad \mathrm{Exp}_A(x, \varnothing_A) = \mu$. In particular, $\mathrm{Ess}_A(\varnothing_A) = \mu$. Such element of the intellectual object **A** we will call its *inexistent*. And then the initial object in the world of intellectual function will be the object $\mathbf{0} = (\{\mu\}, \mathrm{Exp}_{\mathbf{0}}(\mu, \mu) = \mu)$ with the unique for every other object **A** of given category **CΩ-Set** arrow $\mathbf{0} \to \mathbf{A}$ that assigns to the element μ an inexistent $\varnothing_A \in A$.

We will return to inexistent of intellectual object shortly, yet for now we have everything prepared to examine in detail the construction of exponential object $\mathbf{B}^{\mathbf{A}}$, which corresponds to elaboration of the intellectual object **B** by raising it to the power of existent knowledge/representations **A**. Let's not forget that both **A** and **B** are complete Ω-sets right now and that we're working in appropriate category **CΩ-Set**.

What exactly we expect from this kind of object after all? We want it to

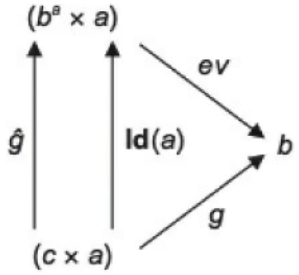
[63] Before we denoted minimal element of lattice Ω as ⊥, or "false but now it is worth emphasizing graphically that it is really minimal possible (actually, nil, or zero) degree of existence.

play a role of a container of sorts filled with signifiers which would effectively name all possible relations of the form $\mathbf{A} \to \mathbf{B}$. That's why we called this construction "quasi-limit since we are really looking inside the world of intellectual function for such an intellectual object that would *effectively* represent all those relations (*parametrize* them "perfectly"). Random choice of object X most likely will fail to do the job in that X may not have appropriate signifier (concept) to express a particular relation, or there may be several different names (signifiers) that express the same relation.

As a simple example of such "suboptimal"object X we can consider a set of buttons of a pocket calculator with the names of functions inscribed on them, for instance: $X = \{+, -, \times, \div, \log, \sqrt{}\}$, and to take the sets of numbers as objects A and B. Then our operations with calculator can be formally reduced to input of a number and a name of an action (function) which we would like to perform, i.e. a pair of the form (a, f) from the set $A \times X$ and getting a value $f(a) = b \in B$ as an output. Obviously, there will be many more relations between sets A and B then there are buttons on any calculator: generally speaking, there are exactly $|B|^{|A|}$ functions from set A to set B - vertical lines here denote that we take into account quantitative aspect of respective sets, i. e. a number of elements in them. This is, by the way, another reason why the metaphor of "raising one object into the power of another"is quite apt in this case. However, it's not enough to know the number of relations between A and B - otherwise we could just enumerate them. But from exponential object we expect the names of relations that are meaningful - we are looking for notions in it that would conceptually grasp those relations and adequately express them. That's why we may say that exponential object contains not only bare names but also the arrows of the form $\mathbf{A} \to \mathbf{B}$ as the result of relations.

Hence, the formal definition is the following: *exponential* (exponential object) of two objects A and B is object B^A together with the *evaluation*

arrow (morphism) $ev : B^A \times A \to B$ such that for every other object C and every arrow $g : C \times A \to B$ there exist unique arrow $\hat{g} : C \to B^A$ which makes the following diagram commute:



It is curious that the arrow $\hat{g}$ is also called the *name* of the arrow g. Object C here plays the role of that "suboptimal"X which lacks some names or, on the contrary, contains equivocal names. The presence of the arrow $\hat{g} : C \to B^A$ informally indicates on symbolic resources powerful just enough to clarify every ambiguity replacing it, so to speak, with *canonical names* contained in B^A.

This is the general case. Yet in the world of intellectual function this object, as might be expected, gains its own additional specific character. Since by world of intellectual function we agreed to understand a category of complete Heyting-valued sets, then on the relations themselves, i.e. on the elements $g \in \mathbf{B^A}$ as well we need to define both the function of expectation and, as a consequence, the measure of their "trustworthines"$\mathrm{Ess}_{\mathbf{B^A}}(g)$. That is why in the category **CΩ-Set** it is convenient to define the exponential object as the set of pairs of the form $\langle g, p\rangle$, where $g : \mathbf{A} \to \mathbf{B}$ is a relation (arrow) in given category and $p \in \Omega$. The function of expectation $\mathrm{Exp}_{\mathbf{B^A}}(\langle g, p\rangle, \langle g', q\rangle)$ is defined so that $\mathrm{Ess}_{\mathbf{B^A}}(\langle g, p\rangle) = p$. This is done in the following, quite specific way:

$$\mathrm{Exp}_{\mathbf{B^A}}(\langle g, p\rangle, \langle g', q\rangle) = \bigcap_{x\in A}(\mathrm{Ess}_A(x) \Rightarrow \mathrm{Exp}_A(g(x), g'(x))) \wedge p \wedge q$$

Note that in this interpretation, only global elements of the set $\mathbf{B^A}$, i. e. ordered pairs of the form $\langle g, \top\rangle$ are true arrows (relations) in **CΩ-Set**, since, according to the definition just introduced, $\mathrm{Ess}_{\mathbf{B^A}}(\langle g, \top\rangle) = \top$ - and it makes sense to consider all other arrows $\langle g, p\rangle$ as some hypostatized bonds which are there (between objects **A** and **B**) only with a certain degree of expectation $p \in \Omega$ that was predicted by the subject of experience Ω. In effect, exponential object, defined in this way, contains both true and *apparent* relations, which perfectly models the work of an intellectual function as an elaboration of intellectual object, where the latter is not always a reconstruction, but could remain just an elaboration of our representation.

Yoneda embedding is devoid of this disadvantage - or, perhaps, we would rather say - of this distinctive feature, because it "forgets"about any arrows in the category except the true ones , by its very construction. Detailed examination of this construction. The last chapter The Yoneda Passion"of our recent work "Language of category theory and "the limits of the world"[64] is devoted to a detailed examination of this construction. For now we will only outline basic principles for constructing Yoneda embedding, making appropriate and necessary emphasis on the existing parallels with the methodology of thought.

So, a functor $\mathrm{Hom}_C(-, B) : C \to \mathbf{Set}$ puts into correspondence to every object A of category C the *set* of arrows from A to B. It means that on the face of it it must be an exponential B^A that is put into correspondence with an object A, but exponential is the object of the category C, while a functor acts into the category of sets. That is why, when constructing the former

[64] *Egorychev I.* Language of category theory and "the limits of the world". SPb. 2018

we had to take into account the internal structure of category C, whereas in this sense there is no any specific structure on **Set**. Actually, we could define exponential also as a functor $(-)^A : C \to C$ which acts from category C into itself (it is called *endofunctor* in such case)). So, to put it differently, we could say that constructing of exponential is "internal"exponentiation, whereas constructing of Yoneda functor is exponentiating "externally".

Please, note that a functor is an analogue of function defined on categories rather then sets. But since, unlike in ordinary sets, there are such "elements"as arrows in category, in order to define functor correctly we must set a rule for their correspondence as well. This rule arises quite naturally: if in category C there is an arrow f between objects A and B then a functor $F : C \to D$ puts it in correspondence with the arrow in category D that goes from and to those objects into which objects A and B were sent by the same functor F (they called *images* of A and B in D under action of F). And of course, such a correspondence must respect the basic category structure, i. e. image of an arrow which is the result of composition of two arrows in category C should be equal with the result of composition of their images in D:

$$F(f \circ g) = F(f) \circ F(g)^{65}$$

To put it more simply, the rule F is *functorial*, if it agrees on arrows.

Correspondingly, there are may be quite many of such functorial rules acting from category C to category **Set** and among them are just mentioned functors of the form: $\mathrm{Hom}_C(-, B) : C \to \mathbf{Set}$. But then, just as in case of exponential object B^A, elements of which in general were arrows $f : A \to B$ we might want to consider another object, elements of which would be

[65]If this icdentity holds, then a rule F is called *covariant* functor. If slightly different identity holds, namely $F(f \circ g) = F(g) \circ F(f)$, then functor F is called *contravariant*. Sometimes they say that contravariant functor "reverses"arrows, since it puts in correspondence with an arrow from X to Y in category C an arrow from $F(Y)$ to $F(X)$ in category D.

functors $F : C \to \mathbf{Set}$. It turns out that such an object really exists and, moreover, it is a category, which by obvious reasons called functor category. It means, in particular, that functors are objects of this category, and that means that there must exist appropriately defined arrows between them. Such arrows between functors are called natural transformations and we examined their construction in detail in our book "Language of category theory and "the limits of the world"mentioned earlier. But now it is quite another aspect that is important. It turns out that from any category C [66] into such functor category $\mathbf{Set}^C$ it's possible to construct a functor which every object $A \in C$ sends to some object in category $\mathbf{Set}^C$, namely, a functor $\mathrm{Hom}_C(-, A)$ which in our case, when $C \subseteq \mathbf{C\Omega\text{-}Set}$, is - attention, please! -a set of *all* relations of intellectual object A with every other intellectual object of category C. And since a relation in a world of intellectual function we defined as ordinary set theoretic function with some additional constrains, a set of all relations between object A and B won't contain anything else except such set theoretic functions which are exactly global elements of exponential object in the world of intellectual function, i. e. relations that truly exist.

Finally, it only remains to add that just mentioned functor $\Im : C \to \mathbf{Set}^C$ is precisely a famous *Yoneda embedding* and it is of interest for us as well because the set of all relations between any two objects A and B in the world of intellectual function turns out to be *naturally isomorphic* to the set of all natural transformations between corresponding functors $\mathrm{Hom}_C(-, A)$ and $\mathrm{Hom}_C(-, B)$. If we denote these hom functors, correspondingly as h_A and h_B, the statement we just made can be written as follows:

[66] To be perfectly accurate we needed to say: from any *locally small* category for a collection of arrows from A to B must constitute a set, since elements of **Set** are sets. (A category is called locally small if, for any pair of objects A and B, the class of morphisms from A to B is a set. Some authors assume this condition as part of the definition of a category.) However, for the purposes of our study such subtleties are not essential.

$$\forall A, B \in C \subseteq \mathbf{C\Omega\text{-}Set} \quad \Im_{A,B} : \mathrm{Hom}_C(A, B) \cong \mathrm{Nat}(h_A, h_B)$$

In terms of A. Kurpatov's methodology of thought it means that reconstruction of the world of intellectual function is really possible.

Let's try to understand what exactly the above natural isomorphism means.

The basic idea consists in that when we substitute each intellectual object with the set of all relations it is engaged in with all other intellectual objects and begin to examine these sets as some autonomous objects, then we, on the one hand, do not loose any information about original objects, and, on the other hand, such shift *in representation* of the world of intellectual function does not import any additional relations or any other types of connections (bonds) into it. But at the same time if before we were limited by uniqueness of representation as by cocoon, now our world of intellectual function turned out to be embedded into a much richer structure which constitutes category $\mathbf{Set}^{\mathbf{C\Omega\text{-}Set}}$ (or it's arbitrary subcategory) - for among intellectual objects A, B, C which belong to the wold of intellectual function there are "other people i. e. new subjects of experience Ω', Ω'', Ω''' etc. In our "Commentaries to "Tractatus"we already pointed out on the deepest connection between the set of parts of $\mathrm{Sub}_C(A)$ any object A and the set of signifiers $\mathrm{Hom}_C(A, \Omega)$ which is expressed with so called *representability* [67] of the set of parts of any object A of category C by unique object of the same category, namely Ω. Since it is true for any Ω whatsoever, it means exactly that functors $h_{\Omega'}, h_{\Omega''}, h_{\Omega'''}$ etc. contain all the information about how, at least at the level of signifiers, *Others* (other subjects of experience) label reality.

In our book "Language of category theory and "the limits of the world"we commented on this very important theoretical result as follows: "... we've got

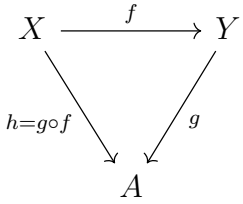
[67] For details see: *Egorychev I.* Language of category theory and "the limits of the world". SPb. 2018, P. 97-98

rigorous and accurate way to represent a certain unknown and isolated part of reality as something familiar..., and second, to immerse it into much more elaborated and also known context which nevertheless have always existed, since $\mathbf{Set}^{C^{op}}$ essentially depends on how the objects of C are related and builds on this information. It means that we have at our disposal a tool with which we can make implicit and unknown relations explicit and to some extent known - and this is extremely important epidemiological result."[68]

One more aspect of the language of category theory that never ceases to marvel, is it's mathematical scrupulous rigor with which the original system of relations can be "dragged"into arbitrary level of abstraction. And now we will attempt to show with all necessary details that such rigor is achieved by following basically the one and the same iterated functoriality rule.

Let's see how this rule is realized using the example of the action of two functors h_A and h_B .

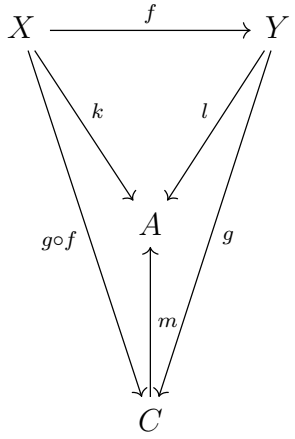
The rule h_A acts from category C to **Set**, putting into correspondence with every object X in C the set of all relations between this object X and fixed object A. Functoriality here would mean that if there exist a relation f between objects X any Y in category C, then this relation should correspond to some other function f^* (dependent on f), which must be a relation between sets of relations $\mathrm{Hom}_C(X, A)$ и $\mathrm{Hom}_C(Y, A)$ in **Set**, such that it would make commute the following obvious triangle of relations:



It is easy to notice that such f^* really exists - every relation $g \in \mathrm{Hom}_C(Y, A)$

[68] *Egorychev I.* Language of category theory and "the limits of the world". SPb. 2018, P. 111-112

it sends to composition $g \circ f \in \mathrm{Hom}_C(X, A)$ (notice also that in process an arrow has been reversed). And if we now consider another triangle:



then we will clearly understand how this rule h_A should act on arrows:

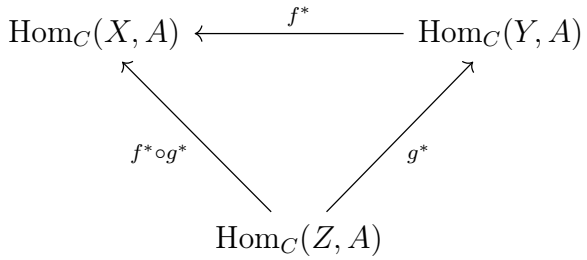
$h_A(f) = f^* : l \to l \circ f = k$

$h_A(g) = g^* : m \to m \circ g = l$

$h_A(g \circ f) = (g \circ f)^* : m \to m \circ (g \circ f) = (m \circ g) \circ f = l \circ f = k$

From the other hand, we see that $f^* \circ g^* : m \to m \circ g \to m \circ g \circ f = k$. Consequently, $(g \circ f)^* = f^* \circ g^*$, which means that $h_A(g \circ f) = h_A(f) \circ h_A(g)$.

That is images of arrows has been reversed indeed, since commuting triangle in **Set** now looks as follows:



And, thereby h_A is contravariant functor.

But Yoneda embedding also acts one more floor up putting into correspondence with every object A in category C just examined functor h_A. Strictly speaking,

situation that has just been examined is called a *contravariant case* of Yoneda embedding and formally should be written slightly different: $\Im : C \to \mathbf{Set}^{C^{op}}$, where C^{op} denotes precisely a category C with reversed arrows (from English word *opposite*)[69].Embedding $\Im$ itself in this case already will be covariant as a functor, which we will now try to demonstrate.

By the way, we could depict functoriality differently. We could say that h_A is contravariant functor if it maps every arrow $f : Y \to X$ of category C^{op} to an arrow $f^* : \mathrm{Hom}_C(X, A) \to \mathrm{Hom}_C(Y, A)$ in such a way that the following diagram commutes:

$$\begin{array}{ccc} Y & \xrightarrow{h_A(Y)} & \mathrm{Hom}_C(Y, A) \\ \Big\downarrow{\scriptstyle f} & & \Big\downarrow{\scriptstyle f^*=g\circ f} \\ X & \xrightarrow[h_A(X)]{} & \mathrm{Hom}_C(X, A) \end{array}$$

Now, the situation with embedding is quite different: we must show that Yoneda functor $\Im$ maps an arrow $h : A \to B$ into an arrow $\alpha : h_A \to h_B$ in such a way that another diagram, built on the images of $\Im$ in category $\mathbf{Set}^{C^{op}}$, commute:

$$\begin{array}{ccccc} X & h_A(X) & \xrightarrow{\alpha_X} & h_B(X) \\ \Big\downarrow{\scriptstyle f} & \Big\uparrow{\scriptstyle f^*} & & \Big\uparrow{\scriptstyle f^*} \\ Y & h_A(Y) & \xrightarrow[\alpha_Y]{} & h_B(Y) \end{array}$$

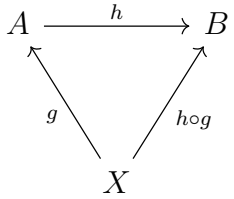
Diagram 1.1

It is important to keep in mind that functor h_A is not just an the object of the category, but the object with the structure which it inherited, as we

[69]That's why exactly this symbol emerged a little earlier in our quote above.

remember, from original category C. In our case given object constitutes a set of sets of relations $\mathrm{Hom}_C(X, A), \mathrm{Hom}_C(Y, A), \mathrm{Hom}_C(Z, A)...$ and relations f^*, g^* etc. between these sets already have been fully agreed upon as well. As soon as this agreement must be maintained on arbitrary high floors of abstraction, the law of correspondence between functors which is called their natural transformation also must satisfy functoriality condition. But now not just one arrow but rather whole family of arrows $\{\alpha_X : \mathrm{Hom}_C(X, A) \to \mathrm{Hom}_C(X, B)\}_{X \in C}$ should correspond to an arrow $h : A \to B$ in such way that for every relation $f : X \to Y \quad \alpha_X \circ f^* = f^* \circ \alpha_Y$.

It's also easy to notice that, just as the last time, such family of arrows (we can denote it as $-_*$) can be found quite "naturally it will be comprised of functions of the form h_*, where each one of them maps every relation $g \in \mathrm{Hom}_C(X, A)$ to relation $h \circ g \in \mathrm{Hom}_C(X, B)$:

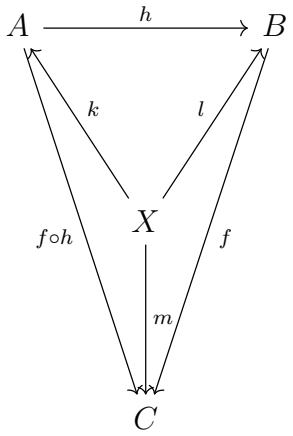


Naturality of this transformation means that the arrangement of every component of given family is *invariant* and do not depend on choice of the object $X \in C$.

If we look again at the identity depicted on the Diagram 1.1 which is sometimes called the *axiom of naturality*[70], then we notice that in our case it is also true that $h_*(f^*(g)) = h \circ (g \circ f) = (h \circ g) \circ f = f^*(h_*(g))$, i. e. the diagram commutes and, consequently, the axiom of naturality holds.

Moreover, it is not too hard to check, that once we have a system of relations:

[70] *Leinster T.* Basic Category Theory. - UK, 1994. P. 28-29



then each component of the transformation $-_*$ is functorial, that is:

$$(f \circ h)_* = f \circ h \circ k = m = f \circ l = f \circ l \circ l = f \circ l \circ h \circ k = f_* \circ h_*$$

So, Yoneda functor maps every object A of category C to the functor $h_A \in \mathbf{Set}^{C^{op}}$ and every relation $h : A \to B$ - to the natural transformation $h_* : h_A \to h_B$ each component of which, in it's turn, is functorial.

Conclusion

The world of intellectual function, realized formally as category of complete Heyting-valued sets allows at least for two more important formal procedures (besides exponentiation and Yoneda embedding that have been just examined) that could also be understood manifestly gnoseologically: the first one implies a radical transformation of function of expectation Exp_A inside the intellectual object $\mathbf{A} := (A, \mathrm{Exp}_A)$. Such transformation can be implemented by joining to the supporting set A the set A itself. This procedure, generally speaking, are forbiden in ZFC axiomatic set theory [71], but ontological consequences of such paradoxical joining are so significant that we believe we have the right to consider them at least theoretically.

The second procedure heavily rest upon the fact that category of complete Heyting-valued sets is a Grothendieck topos and categorically equivalent to category of sheaves over Heyting algebra (over site) Ω. This procedure was once investigated by French philosopher Alain Badiou, and, in particular, the axiom of gluing an arbitrary functor must satisfy in order to be a sheaf[72] was interpreted by Badiou as possibility of a world to be thinkable for the Subject.

Both procedures require a fairly thorough mathematical background, but their potential usefulness for the varieties of implementations of Artificial Intelligence models based on Kurpatov's BPP (Brain Principles Programming) are definitely worth the efforts to gain such background. So, we, for our part, are willing to continue detailed examination both of these procedures and of

[71] The fact that no set can be an element of itself is the one of the consequences of axiom of regularity in ZFC.
See: `https://en.wikipedia.org/wiki/Axiom_of_regularity` for details.
This restriction helped mathematicians, in particular, to eliminate the paradoxes of "naive"set theory, formulated by Georg Cantor, which arose there now and then because of akin autoreferentiality (Russell paradox beeng the most famous among these paradoxes of this class.

[72] See: `https://en.wikipedia.org/wiki/Gluing_axiom`

auxiliary categorial constructions and basic concepts which have been used for constructing them.

References:

Badiou, A. (2009). *Logics of worlds.* New York: Continuum.

Badiou, A. (2014). *Mathematics of the Transcendental.* New York: Bloomsbury.

Baez, J., Stay, M. (2010) *Physics, Topology, Logic and Computation: a Rosetta Stone.* New Structures for Physics (Lecture Notes in Physics 813), B. Coecke (ed.), New York, Springer: 95–172.

Borceux, F. (1994) *Handbook of Categorical Algebra, volume 3,* Cambridge: Cambridge University Press.

Goldblatt R. (2006) *Topoi. The categorial analysis of logic.* New York: Dover Publications Inc.

Egorychev I. (2018) *Language of category theory and "the limits of the world".* SPb: Traktat

Kurpatov A.(2018) *Methodology of thought. A draft.* - SPb: Traktat

Kurpatov A.(2007) *Tractatus psychosophicus.* - Moscow: OLMA Media Group

Lawvere, F. W., Schanuel, S. (1997) *Conceptual Mathematics: A First Introduction to Categories.* Cambridge: Cambridge University Press.

MacLane, S., Moerdijk, I. (1992) *Sheaves in Geometry and Logic.* New York: Springer-Verlag.