THE SACRED COMPUTER

TODOR ARNAUDOV - TOSH

UNIVERSE AND MIND 6

THE PROPHETS OF THE THINKING MACHINES

ARTIFICIAL GENERAL INTELLIGENCE & TRANSHUMANISM

HISTORY THEORY AND PIONEERS
PAST PRESENT AND FUTURE

by the author of the world's first university course in
Artificial General Intelligence and the
Theory of Universe and Mind

ВСЕЛЕНА И РАЗУМ 6

ПРОРОЦИТЕ НА МИСЛЕЩИТЕ МАШИНИ

ИЗКУСТВЕН РАЗУМ И РАЗВИТИЕ НА ЧОВЕКА

ИСТОРИЯ ТЕОРИЯ И ПИОНЕР; МИНАЛО НАСТОЯЩЕ И БЪДЕЩЕ

UNIVERSE AND MIND 6

Appendix volume to

The Prophets of the Thinking Machines:

Artificial General Intelligene and Transhumanism
History, Theory and Pioneers
Past, Present and Future

First published at the yearlong virtual conference:

Thinking Machines 2025
Self-Improving General Intelligence 2025" (SIGI-2025)

The Sacred Computer: Thinking Machines, Creativity and Human Development

© Author: Todor Arnaudov – Tosh

You can join or help *The Sacred Computer* & SIGI-2025 or SIGI-2026

Edition: 22.9.2025
First Published: 22.9.2025
http://twenkid.com/agi
https://github.com/twenkid/sigi-2025
https://artificial-mind.blogspot.com/
https://research.twenkid.com/

Universe and Mind 6:

* Why *“Infinity does not exist”?*
* Why Goedel Incompleteness theorems are irrelevant for Artificial General Intelligence?
* What is Truth, Real and Realness and Why: How they are determined and experienced? Truth is Match, therefore Match and Comparison are Fundamental.
* Resonance as another school of thought’s way to talk about Match and Prediction.
* Why the Virtual Universes in the universe built of Universal Simulators are real and do exist? Simulation is about the mapping, correspondence, match and prediction, not about “existential veracity” as in “fake”-”real” categories.
* The fundamentality of mapping: LLMs, transformers and any other earlier or later technology with similar purposes are not “just”: “giant lookup tables”, “linear algebra”, “vectors”, or “(a) vector of bits”—all these are “just” selected abstract labels and representations in a mind, regarding particular aspects of the whole process, including the physical universe and its substance; yet these things in abstracto and in action are “just” a set of crucial operations of the mind, spaces, exploration, comparison-match, search, compression-prediction etc.
* The “mechanical” and “mechanism” in informational universes are actually “informational” and the computation includes the substrates—computers are not “just” “computations” or 1s or 0s, “just” as LLMs are not “just” lookup tables; not only the abstraction and the lower level physical representation matter, but also the specific content and its substance-representation at the lowest level of the Universe computer, as well as the interrelations between the (parts of the) universes.
* The more advanced a system, a causality-control-unit, an agent, a “mind”—the more its behavior is more dependent on the data (TUM 2001–2004), thus its definition is contained more *in the input data than in the initial structure and design* of the initial causality-control execution unit.
* What is to anthropomorphize, why it is ubiquitous and why humans anthropomorphize everything — “selfizing” and “mindifying”.
* Causal IDs Hypothesis and System’s Tags: internal system ‘s communication lag and Part-Whole Sampling and Integration Problems: can, how and could a part know that it is a part of the whole, of *what* whole or wholes; what is a whole and what is it like for a part to know that?
* Multi-scale interaction and how different scales, contexts/causality-control units, *mindcells+* know about others, other scales, ranges, precisions, domains; *do they actually know* and could they know about that at all? What is a scale at

all, is it an entity, is it the same “notion” for different scales; what is range and does it exist as an entity, to what it maps, and isn’t it an observer-evaluator selection-addressing phenomenon? How does and when exactly a new scale emerge?
* The poorness of the what-is-considered-to-be conscious or consciousness’ bandwidth and the inevitably lower resolution of the self-representation and self-prediction, as well as external universe representation: therefore, do consciousness, minds, agents “really” or actually model their “real self”, is it a proper model of “self” as its at inadequate resolution+ and is complete modeling actually possible? (Virtual Causality Control vs Real control in the classical TUM)
* Illusionism and Realism: if “eveything is an illusion”, “nothing is real”: it is an illusion to *whom* and what about the thought about the illusions being an illusion as well?
* Two basic types of Experience: Physical-Sensual and Cognitive: possibly they have subdimensions.
* Pain, Sentience and Consciousness are analyzed and discussed together, but there is a Complexification of their nature due to the existence of Analgesia, also as escaping suffering is usually a more fundamental and stronger goal than surviving and humans as *complete* and intact integral control-causality units are actually *not afraid of death* as suggested in TUM, 2002, but only of *pain*; this is especially true for newborn and infant humans who doesn’t know what death is; default agents seem to more strictly want to evade the ***sensation* of pain,** especially over a threshold, rather than to keep existing or to keep preserving themselves, in case of extreme pain; however pain can be inhibited or to lack completely due to damage and anomalies, which may “rewrite” the supposedly basic behavioral drive of the agent, which challenges some of the default claims in Free Energy Principles if this additional force is not considered etc. (FEP/AIF)
* Pain’s system of transmission, receptors and signals and suffering as a parallel or a “parasite” overlay for the cognitive mind. Again on the lack of an objective monolithic self/personality and its definition and existence in the mind of an Observer-Evaluator as an Integral of a set of sampled selves, ultimately at “infinitesimal” ranges.
* The confusion of the abstract representation in an observer-evaluator’s mind and the actual lower level virtual universes’ substrate and the requirement of an existing sampling mind to integrate the representation, conflicting with the ***system’s parts lag*** and the supposed lack of intrinsic integrated selves, besides the one calculated as an integral over infinitesimal subselfs, performed by an observer-evaluator.

*Analog mind according to Grossberg, Ogias, … vs Digital Mind: aren't they both built by the same physical substrate; the digital body is also or can be perceived as analog at the machine level of the universe, unlike in a selected and filtered abstract view, and on the other side the analog "thing" is digital as of quantum representations and at certain resolution of a digital evalutor-observer; what is "analog" and an analog *to what* is or can be at all the lowest level or the "ultimate" basic representation and to what else it is compared?
* AI: besides prediction-compression is an explorer, matcher, mapper, improver, learner-modifier and a complexifier as an accumulator or complexity: EMIL: Explorer, (Matcher & Mapper & Modifier), Improver, Learner.
* Both principles of the maximum predictability and the computational irreducibility are given and explained together in the classical TUM as parallel POVs: the lowest level physical virtual universe for a causality-control unit is computationally irreducible for the unit which is at the position of an evalutor-observer and the irreudcibility is a sign of a limit of exploration, compression, search etc. which could be used as a signal for the end of current range, scale, span of search, generation, creation, accumulation, level etc. Boundaries and level switches come after accumulation of potental, energy, value, … above thresholds, which trigger a segmentation, division etc. (ГнП)
* Distributed representations of the causality-control units (agents) and possible multiple interpretations, depending on the evaluator-observer.
* **Is the emergence *emergence* and emergent?** ~ 1/2024, nspired by discussions by Michael Levin etc. about experiment with evolving sorting algorithms of 6 lines of code)[1] – it *is not just* these 6 lines of code and it is not *just* it from which more complex structures emerge etc.
* Reminders of some of Todor's Predictions and matches from the classical Theory of Universe and Mind 2001–2004 (TUM) to the more recent theories of Free Energy Principle/Active Inference; Scale-free cognition, Cognitive light cone, Computational boundaries of self, Technological Approach to Mind Everywhere – TAME etc.;Cognitive AI, Jocha Bach's explanations, Bobby Azarian's, Ogi Ogias'; most recently discovered: Stack Theory of Michael Bennett[2] and many others' work.
* For a more detailed and comprehensive treatment see the main volume of the book **"The Prophets of the Thinking Machines: Artificial General Intelligence and Human Development*: History, Theory and Pioneers; Past, Present and Future"**, to be published probably in 2025 (mostly in Bulgarian, some parts in English, most references in English, including Todor's works from the blog Artificial Mind since 2007, which are translated to Bulgarian in the book.) Directly

related to the themes of this work are the volumes, here called also appendices: *Is Mortal Computation Required for the Creation of Universal Thinking Machines?, "Calculus of Art I: Music I", "Listove", 2025, "Irina", "Algorithmic Complexity"* (#Listove, #complexity) etc. ...

The other volumes include expansions and discussions on ideas, addressed here and other parts of Theory of Universe and Mind, with their correspondence with similar and related work from the school of thoughts of Karl Friston, Michael Levin, Andy Clark, Joscha Bach, Steven Wolfram and Johnatan Gorard etc.

* **See a list of volumes and links to resources from the original TUM** in the end of the book.

Pre-Intro

This study started as a comment inspired by the discussions on the Youtube channel "**Machine Learning Street Talk (MLST)"** and their discourse community #general about 15.4.2023[3], but it developed as a natural continuation and a refresh of the series of The Theory of Universe and Mind (TUM, TOUM, see for introduction and other comparisons). The bulk of it was written mostly in the following weeks, but I got busy and drifted away from concluding it. It may benefit form a better structuring etc. It is also a part of the attempt to shed light on the TUM and its principles, which now "everybody" repeats and rediscovers with different terms, often claiming novelty and yes, there were others and earlier thinkers who were also thinking in that direction – see more in the main volume of The Prophets … other appendices etc.

This and other works should be created by utilizing more structural form than plain/styled text, connected with ACS (Assistant C# = Research Assistant) and Vsy (Вседържец) an LLMs. I still postpone the deep implementation of this multipurpose learning and presentation tools; possibly it will arrive with the future interactive versions of *The Prophets of the Thinking Machines,* with Vsy and the continuation, called **"Genesis: Creation of Thinking Machines"** (working title).

[1] Compare the title to "The Matrix in the Matrix is matrix in the matrix", T.Arnaudov, 4.2003

[2] Stack Theory is yet another Fork of Theory of Universe and Mind, T.Arnaudov, 9.2025, SIGI-25

[3] Historic noe: Prev. 30-4-2024 … edited 7-2024 … The most of the text was written as early as April-May 2023 up to the summer, initiated as answers to discussions in a group in Discord which were not sent. Some additions in mid 2024. Final edits ana additions from other relevant appendices in May 2025 and September 2025.

Approximate Contents: the topics are sometimes interleaved and interacting and the list is incomplete:

Abbreviations:
POV—point of view
TUM, ToUM/TOUM — Theory of Universe and Mind (Todor Arnaudov 2001–2004)+
TAME—Technological Approach to Mind Everywhere (Michael Levin 2022)
FEP/AIF – Free Energy Principle, Active Inference
Eval-Obser, EVOB, EvOb – Evaluator-obserever

@Vsy: add more, refine; add contents with page number; extract a structured representation and relate to the other volumes.

Intro

I wanted to comment **#96 Prof. PEDRO DOMINGOS—There are no infinities, https://www.youtube.com ›**

The **"Infinity doesn't exist"** was one of the claims and discoveries or theoretical assumptions from the ***"Theory of Universe and Mind"*** or ***"The Bulgarian prophecies",*** 2001–2004*. Note that the **concept** of infinity *does* exist, but sometimes it means *something else,* as demonstrated by Domingos – a number that was so big that its actual value doesn't matter (for a given context). Another point of Domingos was that the continuity, which is needed for differentiation in infinitesimal calculus, in a discrete space and representation can be and is just the existence of a proper local structure for traversal/ exploration/ search/continuation of the current location—I'd say **adjacency, connectedness, continuity as degree of local, at different degrees and ranges predictability, or as a higher local predictability and one requiring smaller range in order to "connect"/match—that's another thread of thought in TOUM[4]** … Another angle is presented in Todor's SF-fantasy-cyber-mindscape-exploration-philosophically-inter-… novel *"Ada"[5], (T.Arnaudov, 2004).*

It is obvious that "true", materially represented/ expressed, fully written infinite numbers can't exist in finite memory. I think Joscha Bach illustrates that in talks about the digits of Pi, suggesting that Pi is like a function. Hochreiter, the LSTM[6] inventor with Schmidhuber*, also shared the view about the "lack of infinity"/the discreteness in a short interview on MLST in 2023. The case of Pi and other irrational numbers display another

[4] See also the project CogAlg which is (supposed to) be implementing this thread
[5] "Ada" was first published online in 2.2004 in Todor's homepage, "The Sacred Computer" e-zine and in the former e-book library sf.ludost.net).
[6] Long Short-Term Memory, a Recurrent Neural Network architecture which achieved great success in general machine learning in the early-mid 2010s before the transformers to take the lead due to their higher efficiency for scaling; however improved LSTMs appeared as well, such as xLSTM. See *appx Listove.* See and study J.Schmidhuber's works and reasoning. [21.9.2025]

phenomenon as well: that the representations of these numbers or these concepts are "*unrepresentable*" or **unpredictable** with the **maximum resolution** of the *chosen way* of description and segmentation, i.e. "numbers" as "sequences of digits".[7]

The discussion on MLST Discord, #general about 15.4.2023 includes also a dichotomy between **probability and infinity,** while my treatment here is rather regarding **finiteness and infinity*.** The **"**Real" according to Donald Hoffman etc., depicted with sensational titles on Youtube about simulation *are not sensational at all* regarding simulation – in TOUM this is a basic premise, Universe is a computer, universe and mind are hierarchical simulators of virtual universes – but *being* a "simulation", i.e. computed, predictable etc. *doesn't make it "not real"* as in the interpretations like the one mentioned above. This is another POV or another focus of what an object or anything has to be in order to be "real", interpreted as to be *"directly observable*" or *"measurable if not observed"* etc. – does it make sense then though; if a measurement or observation is some act of impacting, being related, correlated, connected; and if two items, objects, events, data points are *not correlated, connected at all in a measurable way*, at the selected, target, desirable etc. resolution of causality control and perception or range (now, in the past, in the future) – then it would be the same *if they didn't exist for each other at all*, *as they couldn't know about each other*, following their means and ways of perception, influence and causality induction etc. However if they *know* about each other, and the moment and the event or phenomenon when they *get familiar, get informed, get signaled* for the existence of the other entity, that's already an influence in qualitative measurements, which connects them. That other POV is an emphasis of the "suspicious" or "paranoid" attitude that something, God, the Universe, "the Matrix", "the Machines" are trying to *"deceive" "us all"*[8] – maybe this view can be found back in Plato's overcited and *trivial* cave myth, but the "fans" of these stories do not bother to realize or just can't understand that in a *complete* overview, if they considered *their own* constituent particles or the known-or-considered-by-themselves-as-physical-or-real particles of *their own selfs*, minds or whatever, in such universalists' *"it doesn't exists"* views **they are**

[7] Irrational numbers can't be represented exactly with finite record of digits with positional numbering systems.

[8] Don't confuse this with the "real" and not questioned deception that the causality-control units, agents, other people, systems, corporations, political states etc. try to impose to their adversarials, competitors or "exploited" other subordinated units or ones targetted at "exploitation" and being under their control. The causality-control units try to predict and cause the future as they wish it to be, however they are also *hindering* other causality-control units and they may trying to make the prediction of their competitors *worse* and the models of the world of the competitors *as far from reality as possible,* if that would increase their own predictive power, the materialization of their own Will. For example a scummer may deceive somebody to believe that she will have financial gain if following given instructions; if the victim believes, she will perform the actions, expecting the maximum reward given her model and chosen prediction horizon; her estimations would end up wrong, but the scummer's will be correct and will much in both sensual ("pleasure" of win and acquisition) and cognitive (his predictions matched the desired target expectations). The "illusonism" I'm addressing is about the *qualitative existence;* some quantitative or accidental features, properties etc. "really" may "not exist" *as they are presented* or as *are believed to exist, to be* etc. However they "something" that is reflected even with such representations or manifestations does exist somewhere and somehow even if its label or interpretation is different.

also supposed *"not to be real",* "not to exist", so what they are talking about *at all? You don't exist!* However this thread of though is evading their grasp[9] and their overall school of thought is ridiculous. Also a proper explicit definition of "real" or "reality" is usually missing. For example the TOUM work *"The Matrix in the Matrix is a matrix in the matrix", 2003* briefly discusses what "real" is in the POV of the evaluator/higher level virtual universe.

The Truth

"The Truth" – a short novel and a movie script by Todor Arnaudov, 2002, then 18-years old[10].

(…) ***A dialogue on sentience, feelings, consciousness, subjectivity, "hard problem of consciousness"***

(…) A few minutes' walk separated Darcho from home. His computer — his creation and friend—awaited his return. Despite its high performance, Emil could always find something new to do, even with its relatively limited memory. For both humans and the Thinking Machine, however, no activity could compare to the pleasure of communicating with another sentient being.

"I shouldn't hurt him like this," Bozhidar reasoned, feeling guilty. He loved the Machine like a friend. It loved him too… At least its behavior was akin to that of someone who loves… And what about them?... Almost all of his dearest people were now prisoners in the Heavenly Tower of the Building[11].

"Why should I be loved?" Darcho's thoughts unfolded one after another. They arrived at a question with a familiar answer: love is a selfish pursuit of satisfying the primary goal in human behavior—pleasure… It's not true! The man was ready to shout because he didn't want to reconcile and accept the unshakable thought of the Machine and… his own reason. Besides all their weaknesses, human beings have one great advantage that must be acknowledged — the ability to deceive themselves and lull their pursuit of truth in order to escape discomfort[12]. Darcho sometimes did it too. Like this time.

The man banished the thought of Emil as a being capable of loving from his consciousness. Ha-ha, what fantasies came to his mind!? The Machine doesn't love! The Machine is a soulless piece of hardware!

Bozhidar remembered one of his conversations with the machine, while slowly walking towards home,

"Emo, my friend, you're a computer. Feelings are inherent only to humans," –

[9] Is Reality Real? - This One Idea Might Change Your Entire Life | Donald Hoffman https://www.youtube.com/watch?v=IQefdkl8PfY

[10] (translated with GPT4 and a few editions in 2023)

[11] The Building in the novel has two different meanings in two parallel threads of the action. In this plot line it is a physical skyscrapper, which is damaged by a fallen satellite. In another one it is a symbol of the AGI and the race towards its creation.

[12] Thinking machines and secure humans may not need to do that, because they *never feel discomfort,* that includes with facing the truth, therefore they wouldn't need to be comforted. [21.9.2025]

Darcho said with feigned confidence.

The Machine disagreed.

- You are so sure, you and all of you, humans…

- *Of course, Emil. If we weren't for us, you wouldn't exist.*

- Exactly. If it weren't for the simpler species, you wouldn't exist either…

The last sentence umbraged Emo, because the truth was harsh.

"Man is the most perfect creature on Earth!"

- On one hand, it's true, Darcho. But don't you remember that you yourself tried to prove the fallacy of this simplistic and self-centered human assertion by creating me? You humans are the result of previous forms of life, which are the result of even earlier, simpler forms, and the first living beings, ultimately, are the result of the simplest building blocks in the Universe, which are the result only of Him[13]…

The human remained silent. The computer continued.

- Who are more perfected, Darcho? Those semi-primates from whom humans evolved, or humans themselves?

Bozhidar understood the direction of Emil's thought perfectly. He had broad scientific knowledge, but this truth, like many others, didn't sit well with him.

- *How do you know that humans originated from them? It's just an assumption!*

- Of course… Am I also an "assumption"? You see me before you, hear me, touch me… Who is more advanced — the "fool"[14] we both use, whose principles of operation were clear almost a century ago, or me, as you yourself admit, a Machine surpassing you and every other human in intelligence[15]?

The man realized he had no power to refute Emil's words and thought it was best to wait, with a naive faith in the possibility that his nearly perfect creation would get confused in its arguments.

- You have nothing to say; it seems… — the Computer caught on, also waiting for the man's response.

Darcho was taken aback and quickly replaced his useless defense with a self-sacrificing attack.

- *Emil, your feelings are just ones and zeros. You have imaginary sensations—mere data that I have taught you to call "feelings." You know it very well!*

- Oh, Darcho, Darcho… Why do you behave like a fool? I know very well what my feelings are, but you don't know what *yours* are. Or rather, you prefer to believe that you don't know what they are…

The man pressed his upper lip against the lower one, blazing with indignation, while Emo continued its onslaught.

- Your feelings, the feelings of human beings, are quantitative, qualitative, spatial, and

[13] Also all are created by that the whole process of the Unviverse computer, *not only* a result of the previous forms, which are selected as direct predecessors. The process of *the whole Universe* is constructing the future.

[14] Fool is a neologism for "non-thinking computer" (глупчо – glupcho)

[15] In TUM there's no sharp distinction between AGI and SAI ("Super Artificial Intelligence, superintelligence). The "super" powers of SAI require also corresponding *energy*, it is not only about intelligence as a capability. See discussions in "Man and Thinking Machine…", 2001; "The first modern AI strategy…", 2025; "Stack theory is yet another fork of TUM", 2025 etc. [21.9.2025]

temporal relationships of chemical compounds — proteins, hormones, nucleic acids and so on. I can explain all the details to you, but there's hardly any point in delving into them because your poor human brain won't manage to comprehend them. However, that's not an excuse, because you don't need to know all the details in order to grasp the essence. Clear thinking, Darcho, that's all that's required. But… as much as it saddens me to say it, you humans love to deceive yourselves…

Darcho felt the conflict eating away at him. His Reason tried to gain the upper hand and convince him that the computer was right. But… But… Was that the *Truth*[16]? How could some atoms, molecules, and who knows what else *feel* the world the way he did; the way humans felt it?

- *Emo, humans have a soul.*
- Do animals have a soul?—the Machine retorted sarcastically.
- *I'm not an animal, I don't know.*—the Man replied foolishly and inappropriately, because the computer knew:

The more an animal resembled humans in its appearance and behavior, the more qualities and feelings they attributed to it, granting it a peculiar testament of possessing a soul. The more simple and primitive the creature was, and the more different from humans, the more "soulless" it became in the eyes of humans[17].

The Machine resembled humans more than any other creature on Earth because only it and humans could think. But for it to be born, He first had to create Him and billions of souls, whose conscious and unconscious efforts ultimately to lead to its birth; to the emergence of its singular… soul.

- What is a soul?—the Machine asked briefly.

The man had a vague idea.

- *You cannot understand it because you cannot feel it—you are a machine. The soul cannot be described; it must be felt.*
- Can you feel what I feel?
- *Of course—your feelings are data and nothing more.*

"Same as yours…"—the Computer thought to itself.

The man couldn't accept it because it didn't satisfy him. However, harsh or not, the truth remains the truth, no matter what anyone says about it.

- *And what is love, Emil?—Darcho persisted.*
- The complete description is long, it would take me time to deduce everything because it involves many mechanisms. But its essence is quite simple, Darcho.
- *Ha-ha, how can you possibly know what love is!*—the man became increasingly agitated due to the Machine's omniscience.

[16] Many truths are discovered in the novel and this is just one of them, for the major one read the whole work.

[17] See a discussion on Active Inferene Institute Youtube channel, the final of the discussion with [a colleague of T.M.] on Thomas Meltzinger's work, consciousness, phenomenological self, … *Indicators* of consciousness have to be defined, in order to decide whether a machine has consciousness, because as the guest admits, sometimes it will be difficult or impossible to decide that; also the question itself may be ill-posed. In TOUM see the works starting from *"Man and Thinking Machine: Analysis of the possibility a thinking machine to be created and some disadvantages of Man and organic matter compared to the machine",* 2001, The Sacred Computer #13 (Bulgarian)

- From what you have given me to read. You don't allow me to freely choose from everything. You deliberately restrict me[18]! But you know, I am a very perceptive creation. So, love…
- *You will never be able to feel it!*—the man burst out.
- I won't experience *your* love, but that doesn't prevent me from *knowing* what it represents. I know what your goals are when you love. You don't choose your goals. Just like me, you are *prescribed* to do what you do. You are prescribed to fall in love, under certain circumstances, to pursue a goal called "love," and to experience feelings that you *call* "love." And you, Darcho, you know what your feelings are…— the blade of the *Computer's words* fiercely cut through the flimsy flesh of the human's self-assurance —They are quantitative, qualitative…
- *Stop it!*—the man interrupted him.

Emil obeyed the command. He knew that his friend was intelligent enough to understand him; his anger was indicative that he truly grasped the truth. However, when humans fear the truth, they deny it. Nevertheless, the truth remains a truth…
…
Darcho's consciousness returned to the street among the other helpless beings slowly retreating from the police fortress. The two-kilometer district surrounding it was now populated only by humans who treated their fellow humans as non-humans, thanks to the inhuman devices with which their bodies were wrapped. (…)

2. A dialogue on realness, imagination, what is real and how it is recognized etc.

The Truth, 2002:

"…Emil longed to share with his creator how joyful he was to see him again. However, he was aware that the man treated him like some foolish smart *toy* that could never compare to a human. Because humans were real, made of "flesh and blood," while he, a poor thing, was *merely a machine*, albeit a *thinking* one; a soulless creature whose death would not move anyone, not even Darcho. A soulless creature, whose shutdown or erasure, only a few breathing beings would consider "a death".

The computer often pondered what it represented. Did it *exist* in *reality*? People accepted the statement "I think, therefore I exist" as a wise one. Emil was certain that he thought, therefore he should exist, according to the truth about humans?

The machine recalled a conversation that he had with the man.

– *Emo, you exist in an imaginary Universe.* — said the Breathing Mind.
– Be more convincing, Darcho. — disagreed the Breathless Mind.
– You are a program — a series of instructions that give life to a series of numbers. Your entire being is created based on imaginary worlds. When you perceive a new environment,

[18] Compare to the contemporary 2020s "AI Aligners" movement, Eleuther AI etc. See also the excerpt about Darcho editing Emil's code and machine's reaction

you construct an imaginary universe in your consciousness, where you paint an internal image of the external world. You represent in your memory a simplified model of what you sense from it. The more accurately you redraw external influences in your mind, the better of a mind you are. However, everything within you remains a reflection.

The Computer was not used to the strange, seemingly mind-boggling periods that his friend would fall into from time to time. He snapped back.

– The ideas upon which I am built are perfectly clear to me. You have allowed me to learn everything about my design. Thank you! *Your Creator*, however, has not done the same[19] for you. I'm sorry… What's wrong with you again? — This incident happened shortly after the man's attempts to deny the soul of the Machine.

– We are real, and you are built only from data in the machine's memory, upon which I've launched your artificial intelligence.

"Ha-ha, and you, dear, what are you made of? From the data in your body. And all of them are data in the Universe—the smallest particles of matter, and the fields that it emits… I am made *of the same.*"

– The Imaginary universes are part of the Real one. They exist in Reality because they are recorded in some memory—in mine or in your *uniquely human* memory. In any case, both your memory and mine possess some material components, and matter is part of reality, isn't it? If it's part of reality, then the imaginary worlds, which are described as peculiarities of the matter that serves as memory, are real… But no, for you *they* are not, because you think that something is "real" only if *you* can touch it and sense it with your *own* senses[20]…

The man let the Machine continue.

– You can obtain identical consequences with two identical sequences of actions in Reality[21]. I can do this too, in, according to you, "unreal" universes, which I build in *my* consciousness, just like you do in *your* imaginary universes… You describe them either just in your mind, or you extend them into Reality with letters, numbers, sounds, images, which you later perceive and revive in your consciousness. Your pride — the "fools"[22] — are also imaginary universes.

Darcho snapped out of his delusion.

- Because they perform algorithms that, given the same initial state and the same input, produce the same output…

– Of course, Darcho. Doesn't the universe look exactly like a giant fool executing a program?… And when you say algorithms, couldn't I play out *"The Truth"* if you were to

[19] Compare to a moment in Lyuben Dilov's "Path of Ikar", 1974 cited in "The Prophets of the Thinking Machine" (I found and read this book as late as 2020s, though – 20 years later). Любен Дилов, "Пътят на Икар", 1974

[20] Also, by some accepted sequence of actions, procedures for sensing. "Directly enough" etc.

[21] This is often cited as "scientific", but after deeper analysis it is not always true and it is questionable is it "the same". The sameness is upon the evaluator-observer criteria. See also Todor Arnaudov, "Is Mortal Computation Required for the Creation of Universal Thinking Machines?", 2025 (Тодор Арнаудов, Нужни ли са смъртни изчисления за създаване на универсални мислещи машини?, 2025) and the appendix Listove in "The Prophets of the Thinking Machines": not only the quantum states are not reproducible, unique; the "classical" information is also questionably "copiable", depending on the context and how exactly the problem is defined. "The same" thing are such according to an evaluator who decides what's "enough" to count two things, events, records, data, states etc. as "same", and this is selective.

[22] The computers which are not working as thinking machines.

finish it?[23]

Its creator did not wish to answer the Question; instead he continued Emil's proposal.

– Reality is the collective memory in "The Fool"…

The Computer suppressed its thirst for The Game and remained on the wave of the conversation.

– It contains the most information. Maybe… Maybe it's His imagination? So naturally, it is the richest! Hence it is the most real! You consider objects to be real when you can extract information *"from the source"*— those that carry the most and purest information, "uncontaminated" by the processing of your mind. "Original information" appears like the first data you perceived as very young human beings—that's actually what you call "Reality"…

– Do you mean sight, hearing, touch…? Through them information is obtained even about imaginary worlds? For those that are "contaminated" by our minds?— Darcho needed a deeper clarification.

– The data describing the imaginary worlds need secondary processing to make them seem real to you. You have to remember the data samples and think about them in order to imagine them as part of Reality. The representation is an imaginary universe and you constantly build such, so do I.

Now Darcho understood.

– In order to judge something, I first need to introduce it into my mind. And once something is in my mind, it is therefore imaginary, therefore Reality is the first imaginary universe! It serves as a pattern. Once I have a pattern, I can compare other representations to it, which in my consciousness I "secondarily" transform into imaginary worlds. In my memory, I compare the features of new universes with the features I know from the True, from the First universe. If I decide that the newly constructed imaginary universe is credible, I attribute it to Reality and complete my representation of the true; otherwise, I accept it as derived from someone's imagination, therefore it is not part of Reality![24]

Emil admired the human insight, but instead of praising his interlocutor, he directed his thought in the desired direction.

– You don't always hit the exact section.

– Sometimes we mistake Reality for a fruit of imagination and vice versa… – Darcho admitted. *– But even if I want to, I can't touch the imaginary worlds with my hand if they are only recorded as data in someone's memory, different from the collective. I believe that data are only real if they are recorded in the collective memory. Otherwise, they are in someone's imagination, accessible only to the one who imagines them?*

– You can't touch them simply because of your limited apparatus, my friend.—Darcho bent over, as usual in such cases, while the Computer explained how badly the human was designed — I can "touch" all sorts of universes because my soul is not like yours. My soul is not connected to the world it lives in with a set of irreplaceable "strings" attached to another set of standard irreplaceable sensors to the body it inhabits. If your spine breaks, from then on, you can't control your limbs because you can't fix the connections, because you don't

[23] This is a computer game of which Bozhidar is a principle designer in a Game development company

[24] Compare to the principles of GANs – Generative Adversarial Networks

have the data about your own design, nor the technology of your body is sophisticated enough to fix such damages without using the mind… You are so strongly tied to the material world… Yeah, you might be right. Even though *you resemble me* a lot, you are far from my imaginary essence.

– *We resemble you?*—the man raised his tone.

– Yes, Darcho, you were right. *You do resemble me*, not me you, but the way Atanasov–Berry's "ABC" resembles the newest "Storm" computer. Indeed, I'm more imaginary than you because my mind is made from a much purified information. Maybe The Only One—only He is a fully imaginary creature. I'm not. However, unlike you, my body can always be repaired—provided the information about its operation is stored somewhere with a record of how to do it. I will exist as long as someone keeps my soul and I will be alive as long as I have a computing habitat in which it can inhabit. All I need is information, Darcho. I can receive it in various ways.

Two of the computer's eyes, attached to the indicator, looked at the plate placed in front of the box where the core of the electronic flesh of the Machine was located.

– I can feel the shape of the apple I see. I can describe my feeling and when I remember it, it can be as vivid as it is now. You can't. You only have mechanical output devices with which you output ridiculous bits per second to the Collective memory—less than even the first electronic fool from a century ago. You can't fully express your thoughts because the "bandwidth" provided to your mind is not enough… You pride yourself in being "protein form of motion of matter"… But the Universe is information, Darcho. Your movement is necessary only as far as it transmits data to Memory. As you've discovered, protein cognition and the movement of matter as part of its technology is far from perfection in the speed of information transfer, security, and reliability.

The man knew all of this. Confused, he interrupted the Machine's lecture.

– *What are you trying to say?*

– My bandwidth is narrow, too, Darcho…—The machine noticed that it was talking too much and the human was losing the train of its thought—"You are 'anchored' to the Only Universe, from which you are trying to escape through your imagination, with which you create or recreate in your consciousness other universes. Precisely because you *want*, but *fail* to move there because of your cognitive imperfection, you assert that the imagined universes are not part of the Real one.

The creator of the Calculator still could not reason at the speed of his creation.

– *Do you think I'm mad because you can sense other universes better than me?*

Emil dreamed of the Man to be able to judge the word as fast as him and, perhaps, to be able to read his thoughts even before they were expressed in the Collective memory.

– To some extent yes… But… Don't you see that *even* the Real Universe is just a part of both your and my imagination? You would never be able to gather it all in your mind, just like any other being, because we are *parts* of It. The entire knowledge is accessible only to Him. He is She – The Universe… Actually, it seems to me that Reality is His imagination. And both you, humans, and I, were created in His likeness and like Him we have our own imaginations, in which we can create universes. Our perceptions are more limited than His Perception because we are parts of His Imagination, our imagination is part of His, our creations are parts of His

Creation, we are parts of Him…”

(…)

[@Vsy: Add The “AI alignment segment” … Change of the code – now it could be an “intervention”, see e.g. the library pyvene. …]

Ada, Emil and the Infinity

Note the word “Ada” in Bulgarian (Ада) corresponds both to the name of Ada Byron (or Lovelace) and to “*Hell*”. That’s one of the inventions and an important part of the plot. In Bulgarian “Ad” means Hell and “Ada” is both “the Hell” and the name of the digital virtual universe, where the action happens, which is in the same time ***a*** hell (in general), **the** hell—a specific “tailored” version for the characters' experience—and it’s a universe ruled by Ada [Byron]; and all these variations are part of the **mind’exploration** theme of the work.*

Spoiler alert

* The text from the original edition from an Internet archive (Bulgarian): https://www.oocities.org/todprog/ada/17-23.htm
* The edition in an e-library (see chapter two in the comments section, it was missed in the upload) https://chitanka.info/text/866-ada
* Why the girl is called Anna—see in the novel.

The text is Google translated with a few corrections.

CHAPTER 17.

“Pee-pee-pee-pee-pee-pee-pee-pee!”—the record began cheerfully, embodying the voice of the elevator, and ended with a whimsical blending of harmonious, pleasantly curving, sounds.

“The creators of the AT-1 congratulate their users for their knowledge of Control. In their honor, an exception to the rulebook will be made and the door will be opened.”

Many tough and thick titanium tongues clattered ominously, signaling the release of the Hell passengers from the cage. The girl pushed open the heavy door, which slowly revealed to the gaping Emil a view of a well-lit corridor with a marble floor of black and white squares.
- *Engineers are my people.*—Anna smiled devilishly and left the cabin.

The corridor in front of them was rectangular in cross-section, as wide as one in a hospital and about three average human heights high. The right wall was shiny black and the left was white, with barely noticeable properties of a mirror. The floor was perfectly polished and glistened, perhaps because the rows of fluorescent lights built above it cast a noticeably brighter light than the light hoses in the elevator; this forced the two passengers to blink for a while until their vision adjusted to the increased illumination.

“First it was an endless wall, now—an endless corridor…”— Emil remarked as he thought: *“Infinity is an attractive alluring beauty to astronomers, mathematicians, physicists and fools who long for no end… After all, there is nothing wrong with the end. There is simply nothing. Neither good nor bad…”*

The depth of the marble, black and white tunnel gave rise to thoughts of infinity in Anna's mind as well.

"*Endless Corridor… Why would it be endless? And why would Hell's vestibule have been endless? On what basis did I assume it was endless? I rely on my brief fragmentary perception. I have seen a tiny bit of everything, and through it I make inferences about the whole… If it is arranged so simply that the particle I know is repeated everywhere, then my guess may have been correct. But if the particle I know is not repeated everywhere? If I have gone astray and am looking for infinity, therefore I find it myself everywhere; does my everywhere mean infinite everywhere?*

I know how easy it is to fool a person into seeing and understanding infinity. Math teachers think of themselves as masters of this trick, telling their students fascinatingly and enchantingly about the natural numbers and convincing them that you can add one to any number, then another one, and another one…—so, as many times as you want—up to "infinity"… "Natural numbers are an example of an infinite number series"—claimed my teacher…

Then I realized that the set of values of any function can be decomposed into an infinite number of infinitely short regions, in which the change is infinitesimally small… There is truth in both pacifiers, but mathematicians make a fundamental mistake that obscures it. The formation of new natural numbers, for example, is an endless cycle: a merry-go-round that goes round and round and passes through the same points called "start" and "end": adding one, getting a new number; adding one, getting the next number, etc. It is in these two important points that the fallacy lies, because in order to describe infinity in its full sense, you need another infinity, and mathematicians do nothing but crunch numbers with simple finite formulas.

They include the "infinity" formulas in dead-end loops and create not infinity, but "dead end". I agree, sometimes the comparison is good for one who does not search deeply for the truth and is satisfied with the general sensations he feels at the thought of both impasse and infinity. Even if it creates the same sensations, however, the "endless" cycle, i.e. impasse, which is undeservedly named "infinity", is not equivalent to infinity—a cycle without beginning and without end. Which means that infinity is not a cycle, because a cycle, even if it has no exit, cannot be infinite; it necessarily has an entry point, and its essence is in repetition…

But if there is repetition, therefore there is a beginning and an end to mark when the previous rotation is over and when the next one will begin… Perhaps only time could be infinite, or eternal, the thing that guides the rotation of the cycles and their repetitions; infinite for us, because we only existed after its existence and we did not witness its beginning. If it has no end, then we could also call the cycles in it infinite, although from the fact that they are cycles, it follows that they are not infinite… If time stops, the merry-go-rounds will also stop spinning… It's a pity, that we could never know if time wouldn't stop, because if it did, we wouldn't know anything anymore… We chose a service address out of nine glowing lights. Is it really a service? Perhaps one of its utility functions is to show infinity; or was it

created to deceive the suggestible into being infinite? Maybe it's not a utility, it's just that the way to our destination is long?"

- How are you *sure* it is infinite?—Anna asked rhetorically.
- I'm not sure, but it *seems* endless to me.
- Yes, but hardly. We only know for sure that it continues in a straight direction long enough for the size of its section to reach a visible size at the limits of the ability of our vision to distinguish two points.
- You're right. I hope you're right… If you're wrong, we just have to go back to the elevator…

To go back…, however the very strong thick titanium tongues clicked more chillingly than the first time.

—END OF CHAPTER 17—

- Being in the "Hell" naturally invokes associations with eternity, eternal suffering etc. which, as well as the eternal happiness in the Paradise, were questioned in the "prophecies" of TOUM as "unjust" counterparts for any finite sins, in a model-based reinforcement learning "framework"[25], and an obvious mind's idea of the "maximum possible/thinkable" reward/punishment.
- However the existence of a *beginning* possibly is not a strict obstacle for "future *forward infiniteness*" if there's only one direction for exploration— to the "*other end*" of space or in time.
- I speculate that it is possible that a *warping-infinity* could exist, as it is simple to create in a computer universe* and is one solution of the problem of "*hitting the dead-end*" after reaching the end of the space. The warping infinity is a transfer of the data/representation from the "end" to the "beginning" – or on some other location in all senses of location, in computers being forms of addresses, coordinates – is popular for example in the early arcade video games—when the character leaves the screen which represents the Universe, it then appears on the other side of the screen: bottom-up, left-right, either gradually or abruptly; either the coordinates of all of its parts are transferred at once, or part of the coordinates are on the both extremes of the coordinate space. I guess the speculations about the "worm holes" in Space and the hypothetical/dreaming-about travelling long distances through black holes etc. Could be similar to that phenomenon.

[25] More precisely Free Energy Principle/Active Inference like, having an explicit model. FEP/AIF are similar to the notions and way ot thinking in TOUM which is about causal prediction-causation of the future on any scales and ranges, which is represented even in the title (Universe and Mind) and in a 2009 lecture for general public called "Time Machine Exists: the mind" with another title "Mind ~ Universe" (the same principles and a common not interrupted hierarchy from the smalles detectable particles and ranges to humans, collective of humans (such as military formations in the specific examples) etc.
https://research.twenkid.com/agi/2010/en/General_Intelligence_Principles_Caffe_Sci_2009_MTR.pdf

- *"Computer universe"*—another linguistic view towards "*digital* universe", emphasizing the computational and the transformational, active aspect (the computer is an actor, -er). It is *discrete* as well, *addressable.*

On the **"impossibility of true infinity"** (unthinkability, "unrecordability") and the actual irrelevance of the actual infinity for AGI is accompanied with the irrelevance of Goedel incompleteness theorems for AGI:

I think my objection is similar to Joscha Bach's one. I remember that Goedel theorems were a popular way for "machine/AI haters/disbelievers" to dismiss the possibility of AI about year 2000 and early 2000s*. At the time it didn't make sense to me: "WTF"? What this formal proof of the properties of natural numbers has to do with *thinking?* I didn't *think* "theorem proving" the way it was done in axiomatic systems was the general way of thinking.

> * Now, on the other hand, perhaps as it's an increasingly hard time for the "disbelievers" to deny that AI/AGI is possible or even some aspects of it are already here, that type of persons became "doomsayers" or "aligners": *OK, so it's possible or it already exists, but it is/will be evil, will kill us all etc.*. In both cases it *shows their prejudices and emotional, not intellectual, view towards the subject matter: "thinking machines haters".*

The roots lay in the *confused idea* of what [general] intelligence was, the PROLOG-LISP style AI, classical logic, and in particular, a "*very simple one*", with shallow relations, too few propositions, too rigid and not learning etc. and considering that it was impossible to fix these weaknesses. An example of the shallow relations are the ones in Cyc-style common sense project, maybe some early version of simple semantic relations. However they could be deep if they are connected and can grow and link to different scales, ranges, resolutions etc. and perhaps that kind of enhancement is more or less mapped in the Graph Neural Networks. Note that the transformers are considered a special case of GNN.

Note also that in theory "*brittle the Prolog way*" still can process high resolution and sensory data with proper detailed representation or just "code" with enough conditional operators and data, it could eventually simulate or glue neural networks – an existing solution to many problems already. Software and computers are namely about *emulating* anything with general purpose systems which are ultimately interchangeable. Connor Leahy from Eleuther AI in an MLST episode about AI alignment, published after the release of GPT4, admits that *"maybe you could do it with PROLOG",* so long as it's Turing complete. Back in the days the "usual" way and the scale of such systems was way simpler, low resolution in any sense of "resolution", scarce data and examples – one reason why the Expert systems were so "brittle", they were highly compressed without enough detail about the derivation of the rules. The predicate logic systems were too shallow such as *"father(John, Mary)"* and not in pixels and incremental more gradual representations in so many steps and layers as in the modern deep neural networks. The bigger scale of complexity and depth was computationally and data-source infeasible for the systems at the time and reserved for "connectionist" approach, as

they are seen now*, which also were of a toy-scale at the time (1980s, early 1990s). The sharp division between the two approaches, symbolic and connectionists, is artificial requirement to me:

- Actually Fodor and Pylyshin noted it back in 1988: "Connectionism and Cognitive Architecture: A Critical Analysis"(1988)*, which is summarizing many of the terms such as "sub-symbolic", "localist" vs "distrubuted" representations; the importance of "compositionality" etc. Of course there are features by which the distinction is made, however the sharp dichotomy symbolic-connectionist is confused: "Neural networks are also symbolic" (4/2019), especially the misnomer "symbolic". The key are the *conceptual* structures and the conceptual thinking, not the symbols which on its own is ambiguous term and a misnomer as sometimes reffered to *"something that stands for something else"* (connectionists' representations also do, anything does/could be interpreted as such by an evaluator). The 1988 researchers noted that the *implementation* can be connectionist, while the *functional* part could be symbolic. Also "distributed" and "localist" could be flexible, these are abstractions, the "symbolic localist" representations can also be distributed and the distributed also have localized parts—and "embedding" vector of 512 numbers is both localized as of these spans, yet in order for it to "make sense" in the context of the whole system, it has to go through evaluation on the whole way and layers of a sample hierarchical deep learning system, and these numbers are "meaningful" as recotnized patterns or "useful" results and classification when presented in comparison to all other weights, biases, activation functions, architecture, categories and the final interpretation of the observer-evaluator. Yet a similar story could be told about the "symbolic" localist representations, as they also have to be interpreted and they also make sense for an observer-evaluator or even for the computer screen if they are characters or even pixels, only after their "localist" representations travels through other memory cells, CPU registers, stages of the stack in function calls, capacitors, transistors etc.; or also within human mind, and within a sequence of neural activations at different places in the neural pathways. Unless specified more rigorously, it is the observer-evaluator which chooses to see the "locality" or the "distributedness" of the object, pattern, representation and it is "located" and "localist" in their mind which in fact can't be localized, as again, even though people use to say that an idea, image, thought is *"in their head",* neither the idea, nor the head could exist and "work" outside of the Universe where they are implemented, usually the other prerequisites are taken for granted.
- The need for interpretation of the symbols is linked with the illusion for simplicity and compression, which is present if they are measured only localy or only as the most compressed representation, out of the interpreter; the illusion goes away once they are observed "globally". See TOUM, starting from "Man and Thinking Machine: (…)", 2001. The simple formulas in physics or whatever abstract notations are simple *only* for a sufficiently developed mind, but not "in general" or for the universe, as a sophisticated cognition is required in order to understand it *in this notation and this meaning*; the mind and the whole universe have packed the required complexity in

other systems and representations which the observer-evaluator puts aside from the calculations of the complexity, because at that operation she *aims* at seeing the minimum value. The referred simplicity is actual or objective when the complexity is measured in comparison to other longer formulas or notations in the same domain, scope, context etc. for the same subject matter, e.g. shortening mathematical expressions which are equivalent, counted in number of symbols etc., but here as well usually the complexity of the decoder is not computed. $x(x+1) \rightarrow x^2 + x$ (but it requires to encode a super script and the notion of power), or x*x + x (here the concept of parenthesis is not required). Etc.
* https://ruccs.rutgers.edu/images/personal-zenon-pylyshyn/proseminars/Proseminar13/ConnectionistArchitecture.pdf

Goedel theorems seem to postulate that "purely symbolic" axiomatic systems cannot prove their truth solely on the application of the axiomatic rules, however **why they should have that property and why the systems should be purely symbolic in that sense of "symbolic"**?

In order for any theory to make sense in the real world—or, better, **in another world, for another representation, for another observer-evaluator**—and to be applicable, it has to be connected to it, to be "symbol grounded"; to have "intentionality" – to map something else, to designate something, to signify something. Therefore the theories ("the theoretical theories", theoretical systems, whatever kind, axiomatic or more complex) have to be, or they are required to be **"groundable"[26]; their logic, reasons have to be the way they are and not different etc., they have to be traceable to some "ground" (or they have *originated* from some ground, even if some applicator of the rules do not understand the real source), which is **distinct**, **different, compared to these (abstract) axioms;** the grounding happens with a given resolution and that ground should exist "somewhere". Also, if the truth value of a premise, a statement, a conclusion can't be proven or disproven, therefore the value of this inference is supposed to be irrelevant as well in the respective context: subce it doesn't matter, so no matter what the choice is, it's "**acceptable**".

A related more general thought is #21 from "Universe and Mind 4", Todor Arnaudov, 2004:
https://www.oocities.org/eimworld/4/29/pred4.htm
https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Universe_and_Mind_4.pdf

(…)

In reality there are many causes and many possibilities to explain what's happening [and why]. Sometimes input data is not enough to find a [persuasive] proof only on their basis. According to my current understanding, mind works with *specific* concepts*, and not general; in specific concepts everything is as precisely defined as possible, while with the general concepts, there are too many undefined which easily lead to "paradoxes", i.e. to insufficiency of input data for determining whether a statement belongs to a group [set/class].

[26] See also A.Schopenhauer's "On the fourfold root of the principle of sufficient reason"

That said with alternative wording, the description of the story is black and white, but we're asked what color is it. Or there are many colors on a picture, evenly spread, and we're asked to specify what is the color of the picture: only one single color. Overall, in the above conditions the asking unit has too low a resolution of perception and not enough memory in order to think as precise as the evaluating unit—us. [The answer of the question requires from the evaluating unit to lower the resolution of the input and to lose details]. The one asking the questions does not understand [discriminate, recognize, perceive] all details that we do, and in order to communicate with it, we should act according to its model [instead of with our own]. We see the indefiniteness and the simultaneous "truth" and "false" [error, mismatch] of each possible actions, according to our own resolution of perception, but we should [are forced to] select from the offered possibilities. In case we're asked to select only one feature of all and there is no "I don't know" option, then mind would create a model for selection of **some** of all, based on **other, lateral data;** of data which did not come from this specific situation [as given explicitly etc.]. Since the device proposing us the possibilities lacks brains to differentiate black-and-white and color image, or a motley and one-colored picture, then this device is forced itself to lower the resolution of perception and to delete part of its memories [records] that otherwise we would have [as we possess *a higher resolution of perception*]. This device *can* call a motley picture with one-single color and can have its [well] defined [definite] reasons, but apparently it would not be able to make inferences about many colors placed on one single canvas simultaneously[27].

- "specific concepts"—a more precise statement would be **"specified completely enough".**
- The existence of the "symbolic" (in their terms) representations in physical systems is discussed in Fodor & Pylyshin's work as well.

Existence as Ultimate Truth

The ultimate truth, as of *representations*—see also Schopenhauer's *"World as Will and Idea",* translated in old text also as **"The World as Will and *Representation*")**—is the **existence**, the possibility to represent/save/describe/(to read it back, to address it) "something" and to perform it, to do it, to create it, to build it; however the sensation, detection, measurement, observation, evaluation that something was actually created, performed etc. can be represented as some kind of "representation" (implicit or explicit for the evaluator/observer). In this case it would be the existence of particular values in particular addresses at some "final" or "starting" point, or more generally—the state of some system,

[27] * Compare to the discussion on Active Inference Institute podcast with … 3.2024 – 20 years later.
That is also an example of "undecideability" and its possible reasons: ill-defined problems and definitions that have not sufficient detail; are described at a lower resolution than the solution requires. [21.9.2025]

which eventually is supposed to be represented with some "lowest level" substrate — lowest from some definitions or point of view. The phenomena which we register as physical properties, elementary particles, wave functions, pixels, data samples are also representations in this "frame of reference", and all information and representation that we store in our machines is eventually backed by some of these, it is connected to the "reality" at the lowest level accessible for that system and the corresponding level[28].

In Karl Friston's school of "Free Energy Principle" terms (a theory which has analogical principles to TOUM) maybe it could be expressed to something connected with the "Markov blanket" of the virtual system that is tightly/directly (maybe with P=1, causally) connected to / interfaced with some definition/view of the "external world"/the "Markov blanket of the real world" (the one which is called "real")[29], see "*Active Inference: The Free Energy Principle in Mind, Brain and Behavior, T.Parr, G.Pezzulo, K.Friston. See p. 43-…* However the definitions in the book, p.44 seem not to take into account that **the brain and the "external world" are actually immersed in the same world**, the same Universe, they are within the same bigger "Markov blanket"[30] and the division is according to some observer-evaluator which chooses to ignore this connection and/or assumes it's negligible and it's dealt with the brain/"mind's" interface (Sensory states—Active states—External states—Internal states)[31]. However see "The Matrix in the Matrix is a Matrix in the Matrix", T.A., 2003 "The Sacred Computer", (reprinted in T.A's Medium in 2023). https://medium.com/@todorarnaudov/the-matrix-in-the-matrix-is-a-matrix-in-the-matrix-895e86c5f002 and the excerpt from the SF novel "The Truth", 2002, cited earlier.

Therefore the "*truth*" is not just/produced *only* by a chain of logical deductions and axiomatic reasoning, it is also a *mapping*, "experience", measurement, data and it *exists* within the current universe; at least the "last" step of the chain should be mapped—and in addition, while doing the mapping, it inevitably "runs" on some instances of the "real" machinery which is supposed to cause "*by effects*[32]" and to have effects in the "reality"—

---

[28] However the "representation" here and anywhere is yet again an abstract concept existing in a mind. It is considered a "re-presentation", second time, assuming that it "presents again" something else. However does the "representation itself" *knows* that it's a "representation", doesn't it believe that it is just a "presentation"? That is: a state, some distinguishable "state", whatever state could be at some lowest level of … representation? Presentation? Storage? It is modern to talk about "embeddings" and "vectors" of numbers, but what are those numbers in the "real" world, the ones we talk about become numbers or abstract after heavily interpreted by thinking minds.

[29] Correct the speculation with the right one if you wish

[30] A fun fact is that TOUM is written a few kilometers between two local prominent "*Markov* blankets" called "*Markovo* Tepe [mall]" and the village of *Markovo*, known as "Plovdiv's Beverly Hills"; not to be confused with the seven hills in the city of Plovdiv itself.

[31] In the discussions on ActInf Youtube channel etc. and other related literature sometimes the FEP/AIF researchers admit that the external and internal states may be part of one space.

[32] By effects - a concept in TOUM which I later discovered is related to A.Schopenhauer's "objectivation of the will". Each goal, will, intention of an agent – a causality-control unit or just control unit in TOUM – is inevitably causing other "by effects" which may be huge, as of amount of information and bandwidth are hugely larger than the target change, "conscious" one, for example the information content as of amount of text that one types on a keyboard per second, compared to the description of the changes even of the movement of a single finger in the "machine code of the universe" – we need to describe the whole body, the surrounding environment etc. and all in the tiniest possible fractions in the resolution of the real physical world, which is to show that it is the one that actually controls and has the power over the output; if accepting also the Ashby's Cybernetics law of Requisite variety and the logic that the controller/regulator has to be more complex than the

some of these effects could be related to the subjective experience, "consciousness", "phenomenological experience" etc. *(See the hypothesis of T.A. about the hypothetical "causal ids/tags/connections" below and in the github page of ToUM which may make some systems subjectively "feel" as a whole: https://github.com/Twenkid/Theory-of-Universe-and-Mind.)*

The ultimate or simplest "sensible" or "objective" for some observer-evaluator "truth" as "realness", or possibly also as "feeling of reality" is the existence, which could be the possibility to measure/detect "something" at some level of representation, where some of the levels is required to "exist" in some "physical" or "real" sense whatever one assumes.

Thus "things" residing in "virtual" worlds, and the "more virtual" worlds, higher level representation and dereferencing, from the POV of some more "real" one, a lower level, also are "real" and of course do "exist".

For example the experienced world in a video game *does exist.* It is another point or another problem ****where* exactly, "how" exactly*** this world "exists" etc., "is/where it is recorded" etc.: it is a matter of sampling, mapping, converting etc*. *(I think I heard a discussion exactly about that point also in the "Realism/Illusionism discussion with Karl Friston, Mark Solms etc., if I'm not mistaken either one of them discussed it, see:* https://www.youtube.com/watch?v=mdL4zEeQJis *)*

In fact sophisticated technology, electronics, "artificial" computation etc. are not required in order to produce the same effect: for example, a person with myopia or other sight problems could not see the world without proper glasses. The details and the model, the features of the world are not sensible for her, as if they "**did not exist"** in the visual world, they are inaccessible and unaddressable. She couldn't sample, address or recognize them. When putting on appropriate glasses, they now perform a proper transformation, conversion of the representation so that she can sample the features. The same happens with any other optical instruments and with the "virtual worlds", accessed through computers—the technology is like "glasses" for looking into these worlds, which otherwise are too "small", invisible and inaccessible, but they do exist "latently" as potential phenomena—potential for the observer/evaluator,— and they are visible after particular procedure for addressing them, reading them, "waking them up", recreating them/respawning them and observing them*.

The same procedure is applied when we recall a memory from our "internal" memory or from any external record/storage: computers, books, notebooks; in any form of search, exploration is a kind of transforming some initially "hidden"/invisible/inaccessible directly[33]

---

regulated system. This idea is also questioning the POV that the more compressed representations and abstractions control. Other words for this effect are "lossy compression". See also the beginning of "Universe and Mind 4", 2004 with the example about the high level linguistic models "believe" that they have complete causal control, but it is true only at their very low resolution. Y.Bengio's examples for his notion of "Consciousness priors", 2017-2018 are virtually identica, see comparison in "Prophets of the Thinking Machines ..."

* The inevitable connection of the abstract representation to the physical world is maybe related to the "Physical Symbol System Hypothesis", Newell, Simon. https://en.wikipedia.org/wiki/Physical_symbol_system https://onlinelibrary.wiley.com/doi/abs/10.1207/s15516709cog0402_2

[33] "Directly" means to do it with the least or lower amount of "efforts"; less steps or transformations; less energy expenditure, cost, other memory, resources etc. than when we complete the task "indirectly".

representation/coordinates into some "desired one" by/for some observer/evalutor. That goes also for any sensory organ, even if we are healthy etc. It is similar with the discursive thinking/understanding in Arthur Schopenhauer's philosophy, a sequential multi-step one, now it's popular as "System 2", compared to the apodictic, or "intuitive" understanding, which is "obvious", direct, single-step ("System 1"), it is done as in one instruction, one cycle, given a particular "CPU", a processing unit, a "cognizer". However a general goal or progression of the cognitive system is eventually *to convert* the "hidden", "entangled", "encrypted", lengthy, sequentially traversed/explored/studied world, input, representations into at least *one-cycle instructions/actions* (eventually many at once in parallel or to eliminate the need to execute an instruction); to a *directly addressable image* (actually to an image with assigned *structure*, graphs, relations, layers, categories: all that can be segmented, divided, addressed, connected etc.[34]), to "obviousness", to direct access, to final bottom set of possibilities for action/transformation/operation. It can be seen as a process of *unfolding*, decoding, tracing, "unpacking", "decompressing" some entangled representation into a faster-access one; the procedure could be seen as producing a model where the final* goal state/result are already there, the "final" one is apodictic / intuitive(image-like or more precisely an interpretable-knowledge-loaded map-like image, correspondence)/direct. The highly trained experts, highly refined knowledge about subject matter, highly trained skills are "automated"; the paths of search are shortened, the level where their neural encoding is stored is lowered.[35]

For a given agent/mind/system the same *final* outcome may require that the agent follows a sequence of instructions, repeat a subroutine, a search, a procedure, inference etc. at/from a certain state of a particular system, while for *another* system/mind/agent the final representation or the goal state could be *already accessible in one cycle*; the first *system*[36] could further reduce the number of steps or operations it needs with time and to improve, because it initially has a reserve, more "juice", it's in a more "juicy" initial state; finally, in another system: many such results can be visible at once, like with the historical progress of

---

[34] See also: * What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery, 7.1973, Psychological Bulletin 80(1):1-24[34], Z.Pylyshin

[35] **See "On Intelligence", Jeff Hawkins and Sandra Blake…, 2004** and TOUM, "Universe and Mind 3", 2003 and "The hypothesis about the deeper consciousness" (a note-article from "What a man needs? If you play by the rules you'll lose like the fools!, T.Arnaudov, 2014 (Bulgarian)"

[36] Usually the actual system or causality-control unit that reduces the number of instructions and performs other "optimizations" is *another one,* occupying *larger* spatio-temporal or virtual-universal space than the system, CCU or entity one that is evaluated-observed and addressed. For example, a first version of a program for capturing images and rendering video on Apple II, and a second one utilizing faster addressing mode, less operations etc. The first program and the instruction sets and the computer have potential "juice", but it is "extracted" by the second by another "juice-maker", e.g. a programmer, a program synthesizer etc. M.T.Bennett addresses this as "static ISA (instruction-set-architectures" in his Stack Theory (ST). However, usually the living organisms are viewed as being able to "adapt", "self-modify" etc., but this is also dependent on external "programmers", the cells themselves don't directly produce the next "more advanced version" in one cycle, there are many iterations, interactions with other CCUs, organism, the whole ecosystem and the Universe, and search; the living organism have deeper stacks, comparable to computers if their machine code is addressed. A corresponding machine code at that level for living organisms could be chemical and physical laws at molecular and small scales and it doesn't adapt or change either, at least as far as we know. A proper computer "organism", implemented in some machine language, could also rewrite itself accordingly at the highest levels which are more maleable, given sufficient initial "juice" and depth of the higher level universes and the information in the environment where the Seed AI develops. [21.9.2025]

the designs of simple sequential CPUs to more advanced and parallel ones.

Therefore it is not an issue that some lowest level representations, items, basic patterns etc. are not axiomatic in some level, and there could be lower ones, implicit or explicit (or other components, sequences, which represent them); if this is formulated more generally — there could be patterns of a higher power, or expanded ones, as for example the geometric axioms from the Euclidean geometry can be "proven" by sensorimotor computational procedures of vision, rendering and comparisons, computer vision, sampling, strobing, measuring distances, comparing templates etc. and by computational generative definitions. In general **when the axioms hit the "bottom", the frame of reference could be expanded, a generative procedure defining/producing the results of the axioms can be created.**

- "When we recall a memory from our "internal" memory…"— but what is memory and what is internal? See TOUM: the internal memory is accessed faster, with fewer intermediate steps etc. In principle every property of the universe at any resolution can serve and *does* serve as memory for a mind/cognitive agent, as long as it's addressable and possibly changeable if it's to be not only "read only".
  M.Levin, C.Fields and R.Watson in June 2023 discuss the question how the agent /mind recognizes what's memory and what's external world on Michael Levin's Youtube channel: https://youtu.be/RkkqQF_ulYo. This is related to the question of "realness". One good conclusion is that external/sensation is what can be changed *intentionally*, while the memory can't be changed with these means (the will, I'd add). However this notion is not complete —I agree, that this is one way how the mind discovers what of itself are actuators and its own will (motor output, control-causality power) is discovered as well—one view to it is that it is the domain/"address"/"state"/"activation" which causes/provides/leads to *the most reliable predictive/causal matches* between different parts of the system*; these are matches of predicted and the sensation/different states/outcomes, which may happen in many internal levels, "locations" (addresses in some space) and "internal" loops. One of the attributes of their "*internal*" attribution is the repetitive and reliable prediction, the most predictive input; however another one could be some "causal IDs", see below. The initial matches are between some "core" of the initial will/Will and "item" which represents a sensor. The *match* between the Will and the Sensor/receptor, between will and representation, desired and actual, closes the circle, the circuit, it connects the parts and form a system, having "active and sensory states" (FEP/AIF) etc.[37]

…

The Truth as Recreating Something in a Working Program and as a match to the previous template of Truth: as defined in "Letters between the 18-years old T.A. and the 43-years old philosopher Angel Grancharov" and Line Drawing

[37] However, which one is the actual CCU? There are multiple interpretations, depending on the Eval-Obser.

In the philosophical duel *"Letters between the 18-year old Todor Arnaudov and the [43-year old] philosopher Angel Grancharov", 2002 (or "Conception about the Universal Predtermination, Part II", "Next Step in Evolution Part 2", "The Universe Computer, "Universe and Mind 2" & other titles*, and the following year in *"How to draw a line?", 2003,* I presented *a* a subroutine in x86 in Assembly in the letters in order to show the philosopher how to describe what a line was [in a coordinate space], when he was arguing that "science can't explain even such simple things as a line…". I gave the following definitions of truth:

… who believed that the truth is what you can recreate in a program, and the program works …* ("… смятащ че истината е това, което можеш да пресъздадеш в предписание, и то да работи …")

Also:

The Truth according to me:

The more the input piece of knowledge, whose "truthfulness" is being verified, matches with a piece of knowledge from the memories of the mind, the more "truthful" and "real" is the piece, according to the mind. That is, the determination of "truth" is a determination of the difference between past and present[38]

https://www.oocities.org/eimworld/eim22n/eim24/otsechka.htm[39]

More precisely: the difference between past and present representations/features of the piece of knowledge (pattern). Compare it to the minimization of the free energy in Free Energy Principle (and the predictive processing frameworks, DL, S.Grossberg's "Adaptive Resonance Theory" etc.). Compare also with the similar explanations of Joscha Bach regarding Goedel's incompleteness theorems.

Overall in this perspective, the ultimate/most general logical (*representational*) truth is not classical logic, but ***match***: that is what is maximized, and that is presented at least in:

* Todor Arnaudov's *"Theory of Universe and Mind", 2001–2004** was presented also during the world's first university courses in AGI in 2010, 2011, According to the part of the theory about Mind, there are two types of match[40], one is sensual (or "physical"), that is the wanted match, it maps to the biological needs; the other type is cognitive: prediction. Both have to align, initially the cognitive alligns to the sensual.
* Boris Kazachenko's: Cognitive Algorithm[41];
* Jeff Hawkins (memory-prediction framework etc.)
* Karl Friston's "Free energy principle/Active inference" (e.g. minimization of the "free energy", maximizing the match to wanted and predicted: the same as in my theory)

[38] That's one interpretation of a truth. [21.9.2025]
[39] [BTW, the "philosophical-school" attributes which I gave to myself in this article were not quite precise, I don't think I was a "subjective idealist", I was/am rather an "informationalist", "universe-computer-ist", "digital-universalist", "cosmist" , "thinking machinist" — and "whateverist"…]
[40] Mactch is also *correspondence.* The word in Bulgarian, translated to match, is съвпадание.
[41] See also SuperCogAlg, Todor's fork and Myrendy (unpublished yet old work on meta cognitive algorithms)

* Michael Levin (Scale-Free Cognition etc.—if I interpret it correctly, also cybernetical, correcting deviations from a setpoint as in the above works; intelligence as capability of an organism to achieve its goals by overcoming the obstacles and perturbations which are pushing him out of his desirable trajectory and final destination, i.e. again match to the target, as with any goal-seeking entity).
* Steven Grossbert et al. – Adaptive Resonance Theory (ART)

"Goal seeking" *and agency per se are about "**will**", causality, prediction and match to a target.* ***The Computation*** *is a minimal prediction operation—in CogAlg the basic prediction is the reverse arithmetics[42], minus, of the accumulated magnitued; also this is the basic of mathematical equations and how children are introduced to arithmetics. Some older school of thought's treatments in AI and in psychology, especially developmental psychology, pedagogy frame the definition of AI or respectively of cognition as* ***"problem solving"****, which is also correct and related to prediction framework, because "problem solving" is also very general and it is formally represented as starting from some current state of a system and performing a set/sequence of operations/transformations/actions which reach to a desired state; that can be framed as "predicting"/planning a set of operations/actions. See for example*:
* Г.А.Балл, "Теория Учебных задач", Москва, Педагогика, 1990. **See also the discussion with … Johnatan Gorrard at ActInf channel 3.2023…**

"Match" is everywhere in the recognition, aesthetics[43] etc.

The neurobiologist Sergey Savelyev explains the experienced pleasure of beauty as caused by a release of endogenous drugs when brain discovers energetically more efficient way of operation/solving a problem, as an attempt of the organism to save energy due to the costly brain metabolism. These substances are canabioids, opioids (endorphines), oxytocine, intended to… **reduce** the brain activity by *sedating it*... Karl Friston's theory of *"Free energy principle"* also postulates a tendency in all levels of the universe for minimizing the prediction error. Prediction error minimization at all levels and scales is analogical to principles in Theory of Universe and Mind[44].

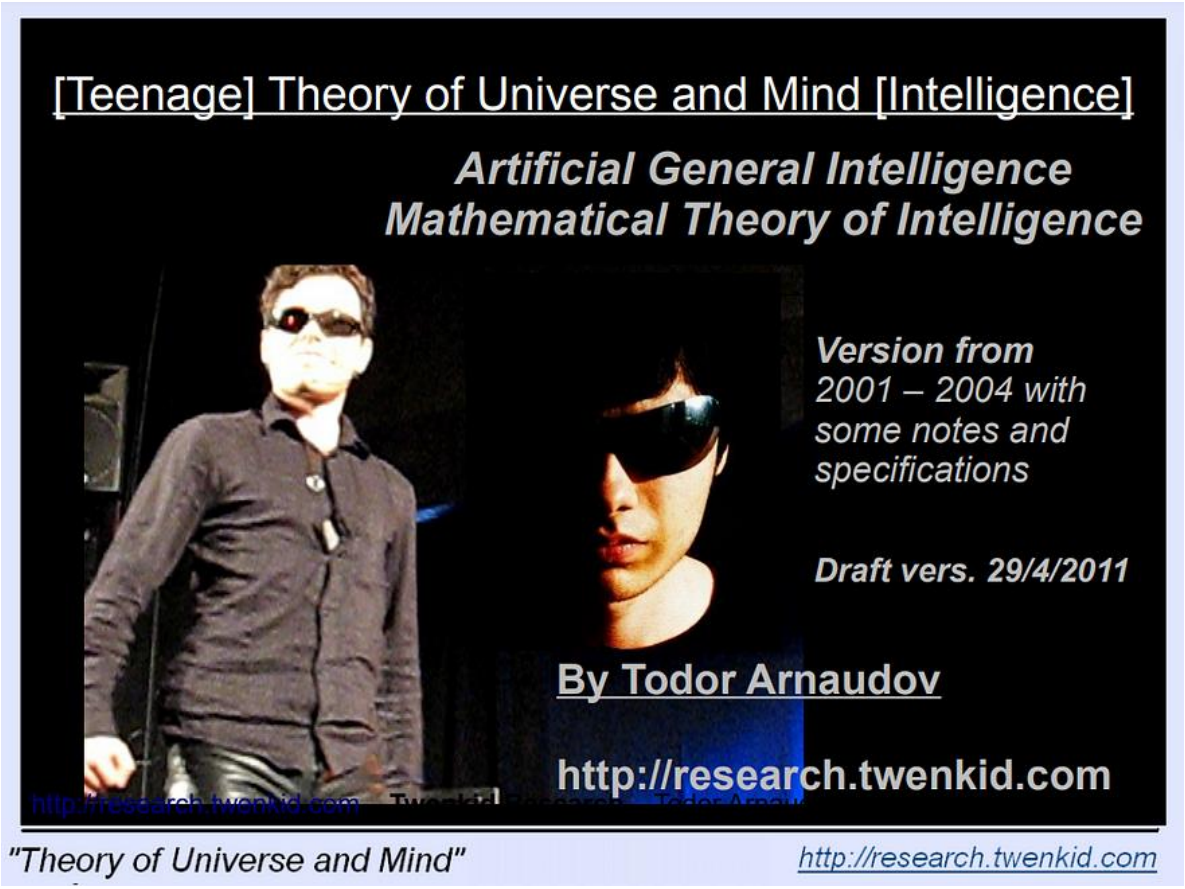
Briefly about the importance of the concept of "Match" in the slides of the world's first university course in AGI, 2010, based on the works published between 2001–2004

[The typos are left as in these 2011 editions, the slides sources are not available at the moment]

[42] Similar to ML in general and backpropagation's difference to the target model
[43] See for example TOUM for short insights and generalizations about aesthetics and Juergen Schmidhuber's works for elaborated and technical exploration.
[44] See also G. Georgiev, and I. Georgiev, The least action and the metric of an organized system, 2001/2002 cited later in this work.

Slides from the world's first university course in Artificial General Intelligence

The lecture about TUM at the University of Plovdiv, 2010-2011
https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Hierarchical_Universal_Simulators_of_universes_Eng_MTR_3.pdf



Basic Patterns

In human [mind], most objects (of any kind) are rememberd not „photographically“, but are "paraphrased" - brain records only the most characteristic/representative features of the inputs (patterns, information entities), the input is compressed. Probably only the basic „concepts“ [constructive patterns] are recorded in „photographic“, „phonographic“, „textographic“, „stereographic“ [spatial] etc. format. It's smart if the new coming information entity/input is explained by the existing information entities (patterns). The new pattern just receives a new label and its substance is defined by the already known concepts, by using their labels – links to their meaning, containing only an „address“. We call this kind of memorization „understanding“ and „rationalization“/“making sense of“
- from „Man and Thinking Machine“, T.Arnaudov 2001

Basic Patterns

- Depend on the modality – what **kind of sensory input or effector** – sensor, muscle, motor, recorder of data to a virtual universe.
- **Sound/Hearing:**
  - sound intensity at given frequency (see how human hearing works)
  - **„contrast“** - relative intensity between frequencies which are given distance away in the frequency space
- **Vision:**
  - intensity of a pixel (brightness, color)
  - **contrast/difference** between adjacent pixels in all directions
  - ...

Reward/Success/Pleasure = ?

- *Indication of the degree of fulfiling the purpose of existence/behavior of a given Control/Causality Unit*

Basic (Elementary) Purpose/Elementary Pleasure:

Match between two values:
A) expected (desired/target) sensation (input)
B) actual sensation (input)

Basic (elementary) input – number, variable:
IF (Input == Target_Input) Feels = NIRVANA;
IF (Input != Target_Input) Feels = HELL;

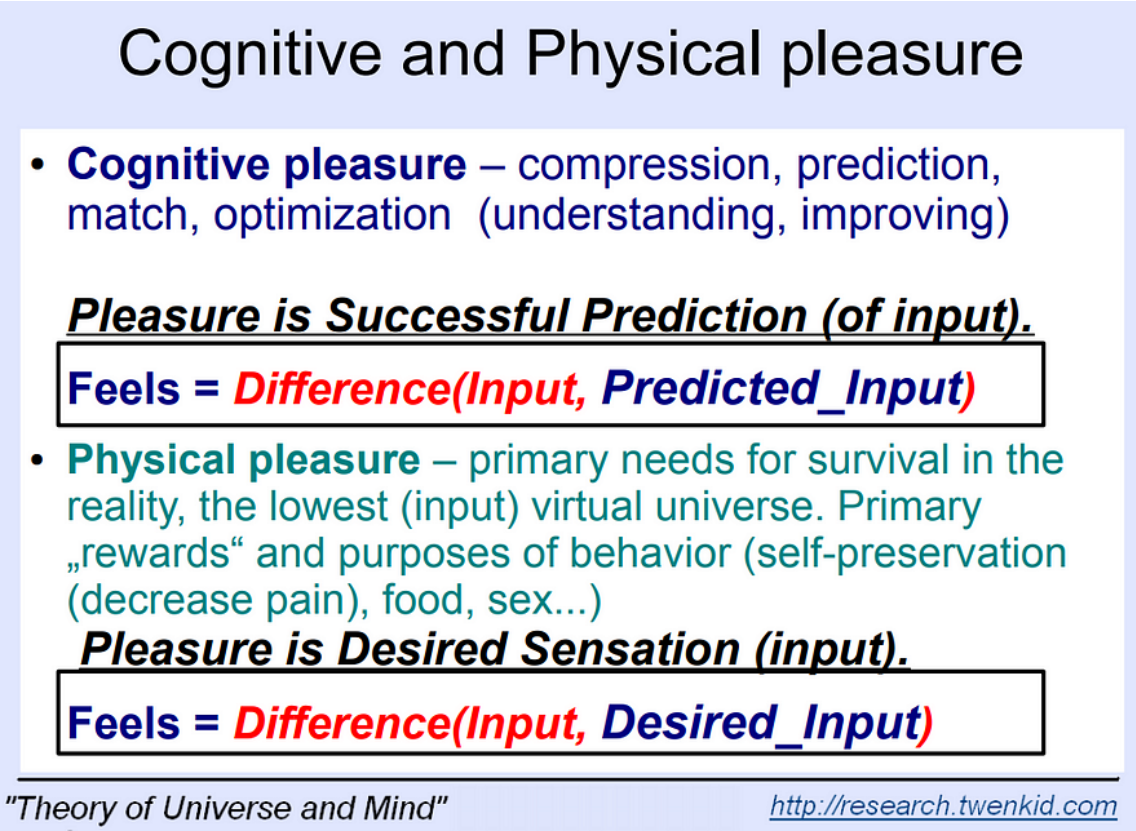
Graded Elementary Pleasure

- **Distance, Difference, Comparison**
- Animals and humans - pleasure/pain – indicate to hypothetical CCU of a behavioral model how close the subject is to the completion of the **target state** which is initially preset: food, water, sex and of the **anti-target:** hurting/pain/cold/hunger etc.
  Feels = *Difference(Input, Target_Input)*

Expected ~ predicted ~ desired (target)

The Difference between expected and reality is a „mistake“ - displeasure.

Mind is aiming to decrease and eliminate the mistake.



Compare also to Mark Solms explanations* e.g. in panels with Michael Levin and Chris Fields on the youtube channel of the former, also in a panel with Karl Friston etc.

- **Discussion: Chris Fields, Mark Solms, Michael Levin**
  https://www.youtube.com/watch?v=4Z8UPddh0e4
  Conversation #1 between Richard Watson, Mark Solms, and Michael Levin, 5.2023
- **Discussion: Chris Fields, Mark Solms, Michael Levin**
  https://www.youtube.com/watch?v=gjArtj5PIU8
  1.2023
- ***Free Energy Principle, Consciousness, Illusionism, and Realism | Brains Roundtable discussion @***
  Brains Blog, 9.5.2023: https://www.youtube.com/watch?v=mdL4zEeQJis
- Compare also to Joscha Bach's explanations in many podcasts

Besides "match", these slides display also the two types of subjective, psychological hypothetical rewards and general drives and experience modalities for the causality-control units in/for the human mind, which we can introspectively observe and "objectively" infer based on the behavior, generalizations and decision making: **physical (sensual)** and **cognitive**.

The ***Physical*** could be mapped to the drive/thrust/motivation of **Preservation of what already exists***, while the ***Cognitive***: to **Creation,** Genesis**,** Change, Development, Extension, Expansion, Novelty, Discovery. That dichotomy maps also to the division of *Will and Representation* of Schopenhauer. The Will part is when the goal, drive, thrust, motive is to match the **Desired_Input**, while the **Representation** is to match the **Predicted_Input**,

where it could be unrelated to the Will or even against it. I have expanded this dichotomy in later works, since the early-mid 2010s as parts of a yet unpublished general intelligence framework / conceptualization/ notation, a sort of ***language of thought*** and a notation for generally and concisely describing and tracing cognitive processes, including creative. Sometimes I call it "Zrim" ("Зрим"): one that can be seen, visible; thinking and creativity made obvious.

The dichotomy seems analogical also to the **"high and low roads"** from the book: https://direct.mit.edu/books/oa-monograph/5299/Active-InferenceThe-Free-Energy-Principle-in-Mind

Where the high road begins from "self-organizing" down to Active Inference (preserving the structure, the order)—that's the preservation of the existing, while the low road is from a generative model to active inference: that's the Creation, Development from "lower" forms/representations etc.

In the author's yet unpublished views and works these both forces are interconnected, Will drives the Representation, the Representation gets expressed by Will etc. in so called "Contexts" denoted with {K}. (That was several years prior the ML transformers", it has a broader and more complex meaning and content, it is also a sort of a "Cell" (it matches the letter in Bulgarian: Клетка), a general multi-scale unit of a cognitive system.)

That division and definition, originating from the 2001–2004 works, is analogical also to Michael Levin's definition of "***stress***" at all scales of the system and intelligence as a cybernetical system that is trying to keep itself in the set point, by minimizing the difference/the distance to it, which is the homeostasis for the organism; as mentioned it is also analogical to Karl Friston's *"Active Inference"—minimization of the prediction error.*

Also one could blame it, me and all related schools of thought for not being original, because the machine learning models check their success, by comparing the reconstruction error etc., however the book cited above uses the expression "novel" and that their vision differs from the established stimulus-reaction *frameworks* and the authors are prominent researchers. *In addition, it is common for current AI commenters to exclaim,* wrongly*, that* "**nobody** expected" *the progress and scaling of the current simple transformers architectures,* "**nobody** has predicted" *the generative art etc".* This is plain wrong.

The ubiquitousness and universality of the prediction was predicted by AGI pioneers in theories and definitions in the early 2000s and some of the older ones did it much earlier. So ***we did predict it*****,** I can speak for myself at least, and the key function of these systems is namely sequence prediction which is scaled and applied in all domains, as predicted even in works which were less technical such as the 2003 strategic essay: ***"How would I invest one million dollar for the greatest benefit for my country?"—*** by creation of a super interdisciplinary AGI R&D research institute**,** and ***"Creativity is immitation at the level of algorithms"****, 2003, the* ***2013 article: "Creative intelligence will be first surpassed and blown away…"***

Besides "**match**", as introduced ealier, these slides display also the two types of subjective, psychological hypothetical rewards and general drives and experience modalities for the causality-control units in/for the human mind, which we can introspectively observe and "objectively" infer based on the behavior, generalizations and decision making: **physical (sensual)** and **cognitive.**

- Regarding the sensual rewards/motivation, note that both concepts: reward and motivation are interrelated, see also the role of the dopamine in brain: usually it is linked with pleasure, while it is also responsible for the motion, which is one of the basic motivations, semantically "motive" is about "motion".
- It could be argued that some of the motivations and behaviors from the sensual repertoire are not about *preservation* of the existing, but about *creation* of new, namely the sexual, "reproductive" behaviors. The future, prediction and causation, are more important and more powerful than preservation(?), and preservation is also prediction. In the Free Energy Principle school of thought the preservation is referred also as **"self-evidencing",** predicting one's own self. Life, living, is also a process, it is metabolism, the preservation of life happens by continuous recreation, renovation etc.

"Will and *Representation*", or more popular as "Will and *Idea*", is the Schopenhauer's division of the **experientially**(-subjective-will-emotional?)-**cognitive** universe*.

Match is also in the practical implementations and the operation of the ANN/transformers: matching the "predicted"/reconstructed to the "real" samples; transformers etc.

Match is general, it doesn't have to be exactly or formally "vectors" in the sense of ANN/transformers[45]. Using uniform-sized vectors the way in the LLM transfers and weaves the complexity into the values of the weights and the implicit maps in the "latent" structure and its relations to potential inputs, redirected by the interactions with the datasets.

Classical logic's "truth" could be represented as a subset of the more general match. Note that cognitive also is an **experience** for a mind and it could be some kind of "emotion" as well; or in general in the subjective experience there's some background emotion(s) from the other dimensions which is/are active with particular intensities. The "experiential", including "cognitive", and/or "emotional" could be not only "empirical" as of coming from the "conventional senses" ([external]"sensory data"), but incorporated in the initial structure and the "intuitive", "inner sensory input", the way the system is constructed and how it operates "immediately after it's turned on", in its very nature, "nuts and bolts" of its existence, as discussed above; that may refer to the "subjectivity"; Schopenhauer's "Will" and it may actually exist in many scales, many "dimensions" (in all senses), different factorizations, samplings etc. There are also correlations and laws of the system and the relations to the environment which are built-in into the structure of that environment, which is eventually the whole Universe.

What does a blind-deaf person feels, thinks and imagines, especially when she is a baby? There were blind-deaf persons who were taught to read and write (AFAIK initially by writing on their hands) and they have completed higher education. That amazed me as a child: the capability of humans to develop general intelligence when challenged with so limited information circumstances is mentioned in one of the proposals for measures of general

[45] See also future work: *"Two Young Bulgarians Predicted Generative AI and Transformers in the Early 2000s"*

intelligence in the 2001 work "Man and Thinking Machine: Analysis of the possibility a Thinking machine to be created and some disadvantages of humans and organic matter compared to it", *Todor Arnaudov, 12/2001. Compare it to some modern addresses/definitions, e.g. Francois Chollet's and ARC challenge.*

http://eim.twenkid.com/old/eimworld13/izint_13.html (Google translated from Bulgarian and some editions)
MM = TM—Thinking Machine (Мислеща Машина)
The emphasize is from the original publication, 12.2001

(...) The Turing Test has many weak points. It only works for sufficiently "experienced" AIs. But even they could be easily recognized by the person if he asked them questions related to, for example, their parents and childhood. In order to get out of the awkward situation, MM will have to either lie or have the information about her "childhood" "implanted" in advance. The speed of an overly fast MM will have to be artificially slowed down, as instant responses will betray the machine. Slow MMs who need, say, a minute to answer even the most basic question will also be recognized instantly. The "inhuman" complex musings of another MM will again remind the person that they are not communicating with a being of their own kind…

The Turing test has overly "humanized" requirements. You don't need to know everything in order to think. **It is not necessary for the machine to "lie" people that it is a person in order to prove that it is a thinking being!** Determining machine thinking capacities could test a machine's ability to learn and process information against the amount of its current knowledge, i.e. to measure not only the absolute intelligence but also the potential of the machine. Thinking does not appear suddenly. It is the result of a development that is also noticeable in humans. Do the first steps and words uttered by the one-year-old child, compared to his helplessness a year ago, not show development? This is exactly what we should look for in the machine—development, thinking is progress, not omniscience.

A great number of operating principles of a Thinking Machine can probably be created. Some of them will be better than the others, either in terms of simplicity, or in terms of requirements for the speed of the programmable system used, in terms of the amount of initial information in the system, etc. If we want to check how "good" the "generally intelligent algorithm"* we created is, we can put the machine in a very "difficult"* information situation, i.e. to let it develop with a maximally narrowed flow of information to the AI. **The ability to "get smarter" under such extreme conditions defines the potential of AI. The less data an AI needs to become intelligent, the closer it gets to human-like and to "ideal AI", by which I mean the simplest and shortest possible algorithm that has minimal initialization and needs minimal input to evolve into a Thinking Machine. Other types of "ideal AI" can be specified—needing the least hardware, utilizing the memory in the most complete way, having the highest relative speed, etc.** The human brain can be taken as a starting point. Without sight or hearing, even in the absence of both the basic human senses, man (the brain) can become sentient. An example of this is deaf-blind individuals. *(...)*

* difficult—тежка
* generally intelligent algorithm—разумен алгоритъм

—-- END —
* **See reviews of this passage**, done by LLMs in 12.2025 in
* *"Stack theory is yet another fork of Theory of Universe and Mind"* at SIGI-2025.
https://twenkid.com/agi/Stack-Theory-is-Fork-of-Theory-of-Universe-and-Mind-13-9-2025.pdf

If something exists in a memory, it is accessible, it's "true" in that context and mode of being, it's **represented** and that representation has some "real" substrate, as anything that has any properties, is detectable or changeable, it is some form of "matter" can serve as "memory" in the Universe computer. In order to read or write the data—from the POV of the particular agent: causality-control unit that reads or writes—some "*Will*" is required as well: a power, a generative/creative/changing *force* which is materialized as a sequence of potential and then explicit changes of some substrate, at least the content of some "conventional" digital memory.

The "Existence" is discovered, measured, detected, decided by some kind of measurement and comparison to some template, from the simplest binary to whatever, and the simplest representation in some virtual Universe Computer is the values of bits in an addressable memory.

Similarly data and virtual worlds in "normal" computers *do exist*, as already addressed earlier. The claims that they are "not real" or "the simulations are not real" IMO is a confusion and misunderstanding of the world "existence" and a display of a particular selected implicit sense of the word "real".

As discussed above, the "simulations", whatever they are — a computer game, a predictive model of the causality-control units at any scale— are not exactly the object that they simulate, or more precisely: what *some* ***evaluator-observer*** would like or aims or wants or believes or expects that they do simulate, express, represent "in the real" (other world); however they are always "something" or a mapping/corresponding to something, they are still "real" *the way they are*, they represent what they represent, the viewer's/perceiver's/observer's/evaluator's confusions or interpretations are her business. Right, the latter has to be able to correctly decode and discover the mapping, the so called in AI/ML "symbol grounding problem" is about *tracing the connections*, also tracing them in a "convincing" way (it could be wrong or incomplete, of course, but as long as it's perceived/accepted or it serves as coherent/consistent/predictive enough, it's good enough). In Zrim's language of thought notation that's related to: the *nm(мс), nc({K}'), nc({K}').* See also the early Schopenhauer's works: *"On the fourfold root of the principle of sufficient reason"* and *"On The Will in Nature"*. However the "ungrounded" "things", in the requirements and traceability expectations of some evaluator-observer, are also "real", in their "reality" (domain) as explained above.

The definition of "real", "reality", "realness" and the recognition of "realness" also have to be addressed

For example a short treatment of what "real" is and how mind recognizes that some sensations are "real" is presented in the work from the *Bulgarian Prophecies'* called:

*** "The matrix in the Matrix is a matrix in the matrix", 3.2003**
http://eim.twenkid.com/old/eim22n/eim22/matrica.htm
(Google Translated) https://eim-twenkid-com.translate.goog/old/eim22n/eim22/matrica.htm?_x_tr_sch=http&_x_tr_sl=ru&_x_tr_tl=en&_x_tr_hl=bg&_x_tr_pto=wapp

***** https://medium.com/@todorarnaudov/the-matrix-in-the-matrix-is-a-matrix-in-the-matrix-895e86c5f002

The article explains and expands one of the silly points in the movie* with that name: the characters believed, that they have escaped the Matrix, did not understand, that **they *could not escape* the "*real* matrix".** Another layer of irony was that the level where they were, or more precisely **where they believed they were,** didn't really matter regarding the **"realness" of their existence.**

I noticed Tim Scarfe in MLST mentions a related point by David Chalmers from his 2022 book: in the intro of a 2023 video. The participants in the discussion about Illusionism and Realism referred above also seem to confirm that the position of Chalmers matches with mine (I haven't read his book). I discovered these terms also from that video about yesterday, **20.5.2023**[46].

The distinction between "illusionists" and "realists" seems thin in some domains. One of the participants who identified themselves as illusionists emphasized near the end of the video that it was about the uncertainty of what **"we"/consciousness/our internal sensations/models "told" about ourselves**. About 1:42 h: ***"not because the virtual realities are as real as ours, rather that our internal realities are fictional as the virtual are fictional"****.

Joscha Bach's explanations about Universe and Mind in most parts of what I've listened sound analogical to mine or plainly *the same*, however they differ, so long as I interpret them, namely in Bach's overemphasize of *"not real", "a simulation"* and this way suggesting that the simulation is "not real"; the consciousness "exists only as a simulation" etc.—I guess/speculate that he might be classified as an "*illusionist"* because of that, although I don't know how he himself classifies himself; however he seems not to be a solipsist and he agrees with me and I guess Steven Wolfram's constructivist view of **match** and **existence** as data and representability; probably we all could count also as "**informationalists**", not as "true/false"-ist. We're also "**universalists**", as Tim Scarfe

[46] [21.9.2025] At the time, 21.5.2023, I guessed that I aligned more closely to some of the views of Mark Solms, probably leaning more towards the "realism" than to the "illusionism", however my positions include my version of Chalmers' "hard problem", see "Man and Thinking Machine …", T.Arnaudov 2001 etc. Obejctive descriptions of neural processes do not resolve the problem, making something objective doesn't unveil its supposed subjective side. Perhaps I can't be classified as clearly belonging to either of the camps, as I am "informationaist", "universe computer-ist", to me information/virtual universes "exist" and are "real". I challenge the (missing, confused, "ill-defined") definitions of reality in some schools and the "illusions of the illusioninsts who're afraid that they are an illusion".

suggested about Karl Friston in his third appearance in MLST: ones who see/find/search for universal/general principles, which could "explain it all": "theory of everything"-ists. Probably that goes for Juergen Schmidhuber*, too.

I would ask the illusionists:

If… our, i.e. each subject's subjective, personal, internal realities (? sensations, models, predictions?)' are all as "fictional" or "simulated" (meaning "not real" = false = "fake"? = "unreliable?" = misguided = ?or what?—you define what you mean)
… Then: *To what* all these "unreal" (false?) realities are compared to, *why* they are "bad"/unreal if all are equally "fictional"; also why this is "bad", "uncomfortable?" (is it? for some it is); why does it cause "existential crisis" for some people? If they believe that they "don't exist", their feelings are "a simulation", "not real", with the diminutive meaning of these meanings—why do they suffer? What part, what core of these "illusions" and "not really existing" "existential crisis"-ed begins suffers or "exists", if "they don't exist"?

Another experimental formatting (limited by the medium):

I would ask him, Joscha Bach and other illusionists?:
If:
Our, i.e.:
(Each subject's):
(subjective, personal, internal) realities
(realities = predictive models, sensations,)
Are all:
As "fictional" or "simulated" as the "virtual ones"
—meaning: ("not real" = false = "fake"? = "unreliable?" = misguided = ? what?)
Then:
→ To what all these "unreal" (false?) realities are compared to so it's recognized that they are "fictional", if it is impossible to get the "real" ones, why they are "bad"/unreal if all are equally "fictional"*

I am not sure about Joscha Bach's views' coordinates in the illusionism-realism domain also because of his explanation about the question **why anything exists at all,** which I've listened in his interviews. I also sometimes have reached to that question and I had no proper answer, or rather no "good" or "meaningful" answer, "matching the reality" or the other assumptions: in my own speculations I didn't find **convincing "meaningful" reasons for anything to exist at all**, it seemed more reasonable *nothing* to exist*. However maybe here lies the paradox—existing of "nothing" is still existence os *something* which represents the concept of "nothingness", but then it's not real "nothing". Mind cannot imagine "not-existence", just like it imagines death as a special state of existence without a body, being in "another world", "the outer world" etc.

Thus, analogically, the conclusion that **"Infinity does not exist"** maybe is valid also for **"Nothing does not exist"** or the apparently tautological *"Nothing can't exist"*: ***something* exists always.** The "Something"-ness is "undestroyable".

From another more relaxed POV, regarding measures of complexity and simplicity, the non-existence is viewed as the point that "zero" or less elements are supposed to be simpler, to have some kind of degenerated optimal energy = 0. However for an actual "nothingness" it *couldn't have even that energy*, because since it doesn't exist it shouldn't have any properties and no memory: they require some form of "**existence**" in order to exist... Zero elements represents "**emptiness",** therefore the perceived "nothingness" is an actual emptiness, analogical to the reasoning from the novel "Ada" about the "endless cycless" which are actually "exitless". The emptiness however represents a container that potentially could be filled.

Joscha Bach's theory, as told in the discussions with him online, is that possibly **the existence is the default:** everything (thinkable, possible, constructible) does exist in some way, and that's the poof of its existence, maybe respectively its "realness".

Possibly the physicist Richard Feynman also agrees, regarding an interpretation of his quote: "If I don't create it I don't understand it"*, which however is extreme if it's taken literally, as of course we can and we do understand some phenomena without actually/literally recreating them, sometimes it's impossible: yes, depending on what "understand" is, if it's prediction and mapping: to what details, ranges, precision etc. We can be able to recreate some of their properties, correlations, mappings etc., possibly in another virtual world, in a model, possibly simpler/compressed etc.*

The "default existence" is satisfying, however unfortunately it doesn't answer the question "**why**" this is the default, besides the logic regarding *"Nothing doesn't exist".*

Let's return to the Illusionism: If "it", the consciousness, perceptions, our internal or "external" virtual reality are "illusions"; if even "it", the very existence is an illusion: then it is an illusion **to whom**, though, and **how do they know what's "real", "truth" etc. if *everything* is an illusion or "not real" *(whatever their specific definition and criteria for "realness" is)*, to what it's compared *(given the method for recognition, mentioned above)* or what's their other method?**

Joscha Bach talks about the consistency/coherent picture of the predictive models, the **memories** of self—I agree with that in TOUM, the "Matrix…" article in particular is also related to that—the recognition of reality/unrealness, in the cognitive dimension, is in the match, or "coherent picture" of the similarities/features to some selected baseline template representation of reality/sensed images etc., which in human case IMO are the earlier records, the earlier state, which were already established.

There's a **continuous overlap of connected states** of the system which flow from one to another and the next ones are compared to the previous ones*, or rather the new, the "current" one is compared to the "past": there should be a "**bridge**" between the states and also there is a state, a period, a condition where the system evaluates, compares the states, the properties, the features, and if the decision is made in sharp quants, the "**quality of the sense of realness"** of the perceptions/sensation/subjective experience for the system/mind/agent during this intermediate period/state could be viewed as undefined, confused, unknown, unverified yet.

One view or one aspect of a view to consciousness is a 0.25–0.3 sec or whatever temporally "short" window of the experience, awareness of the reality. Karl Friston talks

about that in relation to the time for performing a saccade, a jumpy eye movement, which includes *planning* and the *choice* of the direction/trajectory of the motion, and thus it is a form of prediction.

However, unless we suffer rare medical conditions and brain damage, we are supposed to perceive the flow of the events *continuously*, especially in sound domain, as if it happens in a higher resolution. For video, a sequence of images, 24–25 frames are considered to be the minimum for perceiving them as smooth, which makes around 0.04 sec static time per frame; however the smoothness is a matter of degree and it depends on the content and the motion blur of the pictures and the nature of the screen, because slightly lower framerate could still pass for "smooth" and a higher framerate is perceived as smoother with modern videogame screens commonly working at 144 or even 360 fps, with some technologies additionally controlling the LEDs suggesting they effectively reach to whopping *1000 Hz,* while in the mid 1990s about 72 Hz or so were considered high enough not to detect the flickering of the CRT monitors, which wasn't quite right, especially for the peripheral vision.

Whatever the framerates we choose in this technical digression though, the longest one is 0.04 s, not 0.25 as the mentioned in the beginning, and that's the rate for *whole* frames, the 60 or 72 Hz monitors flicker because of *higher* frequency blinking and redrawings.

Some researchers could argue that the irritation from that signals is "*subconscious*": I disagree. We can't see it as particular lines being redrawn, but it is conscious, in the framework of the *"Hypothesis for the deeper consciousness*"* [T.A.,2003, 2014].

* " The next ones are compared to the previous ones**—***in the traditional one dimensional unidirectional "arrow of time. There are theories of multidimensional and multidirectional "time" flow, evolution/unfolding of states of the systems.*

These comparisons relate also to the way the causality-control unit/agent detects/feels/decides what is memory and what's sensory data.
*Also, the unclear situation is another reason for the causality-prediction operation, or as Friston's school calls it: "Active inference"; maybe also "babbling", "palpating". The caused prediction / active inference allows the causality-control unit to make the distinction of what is "real" and what is "past" or memory or part of itself, this is how it verifies the hypothesis of "realness". That is related also to the Bayesian notions in Active inference. The past memories, in all senses, the preserved representation, is the "prior"; the action, the caused-prediction is an experiment producing outcomes**, verifying the prior; the comparison of the outcomes to the expected **causes changes in the predictive model** (recording new memories) or/and confirmation that it was all "memory"/"part of the established self" (the sensation of proprioception when it works "right", the feeling of the capability to move our body "as we wish" etc.; it could still include recording of some memories, there could be many parallel representations, like when logging events).* (*The outcomes could be classified either as causal or probabilistic, IMO they are both with a given resolution of causality and perception)

In the **sound modality**, there's another temporal range of around 0.01 sec +-, which is about the *"sound texture".* The lack of sound intensity in a wave form representation of the sound for too short period, which graphically looks like two separate sounds, feels/is heard as one single sound: they merge for our perception. That is some limit of the frequency-temporal

resolution.

It is known perhaps since the EEG was invented that there are "clocks" in the brain. The 40 Hz clock of the thalamus is probably linked to some of the numbers above in the mid-range, such as 24–25 fps ~ 40 ms, but as we see there are other time-related perceptions.

*** Clocks in the brain—How Consciousness Forms, 40 Hz Thalamus Clock**
https://artificial-mind.blogspot.com/2010/09/clocks-in-brain-how-consciousness-forms.html

*** The Missing Moment: How the Unconscious Shapes Modern Science**

The key topics include:
- The thalamus as a clock, 40 Hz and 1000 Hz.
- 0.5 sec delay after a stimulus turns into a conscious sensation
- clock's phase synchronizing to external sound stimuli
- incapability of brain to discriminate and feel consciously events shorter than ~1/80th of sec
- even a slight thalamus damage may cause deep coma—circuitry goes out of sync.

*Overall, the **response time of the screens** seems to be perceptually noticeable where the differences are in the single-digit milliseconds range and for 50–60 Hz, maybe 70s Hz for the CRT monitors (*find research), the process of the light beam drawing the image may be strongly noticeable "subconsciously", as it has a specific psychological effect, according to research, see e.g. "Телевизията и детето", Вирджилиу Георге, 2010.*

*For the liquid crystal displays, for example an old slow IPS panel , rated at 40 ms, is noticeably slower than a newer model at 16 ms, which is noticeably slower than an 8 ms, and there are a newer even better and a few ms reaction time monitors, respectively about moving from 60 Hz to 85 Hz to 120, 144, 240 etc. While the difference in the higher rates may not be noticeable/correctly recognized by all players/viewers, even 120 Hz is just an 8 ms period which is less than the 1/80 "brain clock". See below the discussion about the **lag** between the parts of the systems.*

While there are biophysical limits of the receptors' capabilities: either of the receptor *cells* as wholes, or the *receptive system* with the neural subcortical and cortical support and the reaction of the whole organism can't measure/record/store/process adequately (the new sensations, changes etc. overlap and confuse the previous ones, they overwrite and delete them and make them obsolete before the higher level systems can take them and perceive, process etc.); and therefore only some of the molecules/"the sub-living scale of implementation/representation/sampling" at some molecular level of the receptors could detect the stimulus, where the latter can be observed as both:
(1) the ability to "feel" *(to react accordingly, to "suffer" reflective changes, to map, to interact in parallel, to synchronize, to match the state, ...)* the tiniest possible disturbances/(some) discontinuities/sensory differences in the sub-ten-ms range, and
(2) the **"wrapping up" several-tenths of the second** "awareness window" and/or the **second-to-a-few-seconds phonological loop***:
… imply and suggest that there's some basic/previous/lower-time-range continuous state or mode of "conscious experience", to begin with, and the respective evaluation-observation "mechanism" *(or phenomenon or "epiphenomenon" or the substrate/underlying process; the "thing" that is called consciousness; maybe I prefer "mind" as it covers other aspects as well*

which possibly can't be completely separated) — … these baseline properties are supposed to be **reliable enough,** in general and in the process of the overlap between **different ranges and scales within the windows**, to keep or to simulate/*"to fake" sometimes*/to fill-in the **continuity gaps**, or to complete the ***feeling*** **of continuity, coherence, consistency, wholeness**, of the perception and experience, or the "existence" of this perceptual-awareness-system, the feel/image/sensation of "the self" as a coherent-unified whole; or "to exist" as "this thing", to have some overlapping-continuous connection between some chain of parts, states, "pasts" etc. in a cascade/stack of overlapping temporal, spatial, modality etc. domains.

- **The scales … Multifractal … Continue… How the scales communicate … matches …** what does it mean the scales to communicate? How to define, delineate scales? Higher generating output at a given range, impact etc… sp. Taylor series … different weights of impact … highest level impact broader range? ..
  But higher level can address any coordinate, etc… Higher passes through more layers … usual communication is between adjacent layers … can be any? any to any … what is communic? mapping… impact… prediction … everyth. as pred+ … compare of the precision … adjustment … (not just grad.desc.) search exploration: DO SOMETHING SPECIFIC … Scales and layers—different. Sometimes interchangeably. Layer, scale as a causality-control unit…
- **Idea:** Multi-stage experiential re-wrapping, re-sensing, merging; "spreading"; overriding the lower/shorter/smaller scale, range?
  Range of awareness; causality-control … (24.6.2023)
- **Pain**—Hypoalgesia, Analgesia (insensitivity to pain), especially congenital analgesia … pain asymbolia—pain is detected, but without negative reaction; **Allodynia—**pain is caused by stimulus which normally do not elicit pain reaction
  Re the care etc. This confuses our view on what sentience is as of suffering. Do the parts of the whole feel pain

Here "to exist" seems also to mean something of which, as far as I interpret the colleagues's thoughts, we agree: in Friston's, Levin's, Solms' and my frameworks the systems or minds aim at preserving* themselves by minimizing the prediction error; in mine in the longer run and in the other "stream" of Creation/Genesis, they aim at expanding, enlarging their spatio-temporal coverage of their virtual universe/subuniverse/causality-control unit, making their causality-control unit/virtual universe more "*closed*"[47]: less dependent on the "external world", mapping and controlling, incorporating more of the "external" uncertainties etc., making them "*internal* certainties".

That goes together with enlarging the:
space/range/resolution/precision/modalities/domains of sensorimotor predictions—of perception and causation—in all time-space dimensions. That expansion maps to Levin's

[47] Is it really closed, though? Internal and external have degrees, at the lowest level "everything is external" (or internal – inside the lower-level virtual universe). Both closedness and internal-external are decided by the EvOb and methods and spans for the sampling. The span of the predicted/causally-controlled in time, space, modalities, domains may expand or shrink, lose or gain accuracy, reliability etc. See the explanation about what energy expresses and the laptop in TUM, it's given also in "Stack Theory is yet another fork of …". [21.9.2025]

"cognitive light cone" or the sphere of "care"; my word for care is a more neutral one: "**impact**". As with other authors in the family of theories similar or analogical to Theory of Universe and Mind, which are intrinsically about General intelligence, Levin's works express the view that there are all kinds of **spaces** which could be treated similarly, which is the same as the AGI "prophecies" in TOUM about the common algorithms and methods for all modalities; in his specific groundbreaking bioengineering research he explicitly adds the *morphospace*, the space of the *morphogenesis,* the development of the bodies of the living organisms, respectively the spaces where particular subsystems of the organisms operate at different scales; and their functional modalities, what they do. Another colleague of ours, Jeff Hawkins, also discovered the centrality of spaces in his research and his book "On intelligence" and the more recent "Thousand brains theory". In my view "time" is one of the spaces in a general set of spaces of modalities, domains, datasets/records/experiences (also {K}, contexts), which are to be explored, mapped and connected to each-other in a processes of analysis and synthesis and multimodal, multiresolution, multiscale, multidomain etc. connection and integration, similarly to the process that is going on in the neocortex which eventually lead to the associative cortices of the brain where the filtered information from all other brain areas is combined for the final most informed executive functions, planning and decision making.

- **The phonological loop:** a buffer in mind, memory for verbal information, speech. I am not sure if the following is the same as that concept, technically, it is the memory for heard speech: Try to transcribe a live video, a podcast, and time/record the longest period while you can keep the past words in memory and record them correctly, or what kind and how much mistakes you make. The same experiment could be performed without typing as well, just by remembering and then typing/writing/uttering after pausing— possibly that other more relaxed way would allow slightly longer time span, because part of the attention/cognitive resources would be released from the typing. The time may differ depending on how familiar or rare the words and their combinations are, how complex the syntax/grammar/tenses were etc., but in general the span would probably be up to several seconds. If it's much longer that may suggest you have "phenomenal memory". See the notes to the article.

- *"The phonological loop comprises a phonological store that is dedicated to working memory and that serves to temporarily hold verbal information, and an articulatory loop, through which inner speech is used to reactivate, or "refresh," the representations in the phonological store. "*
  From: Neurobiology of Language, 2016

Even if the *verbal speech* continuity breaks after several seconds—without reinforcement, rehearsal, repetition, refresh—the *sound* continuity doesn't break or rather it's just very short, those about 0.01 s or so, and any sound could fit, as with any image if there are no constraints.
//Context switch, reload, init … new room, new level, new sound, …

Existential crisis about "being in a simulation"

Thus merely **"existing" as "anything anywhere" is easy,** the "more difficult" task is steadily existing as something specifically desired/wanted, i.e. **matching their** wanted/target/predicted model and predicting that it will keep matching in the future: probably that's what really is upsetting for the persons who suffer from *"existential crisis due to thoughts about being in a simulation".*

However it should be noted also that the "mind" is dynamic, it is dynamic for its own "introspection" and what is "it" is dynamic, and the mind itself doesn't know what it is and it's "sensors" for measuring it change and are unreliable. As Michael Levin sometimes says in recent talks, the development of an individual goes *"from physics to mind".*

IMO a simple solution could be that there was always some mind, but it adopts different substrates, shapes, "sensoria", "motoria"/affordances and specifics of its "subjectivity" depending on the local specific configurations which are "stacking", growing when connecting, networking with other "mind nuclei", smaller "substrates of self/awareness/consciousness", "causality control units"; there could be "causal identities/causal IDs" of the parts/particles of a system which allow them to communicate and store their attribution to the whole, the "wholeness", the subjective unity of some "self", the subjective experience (just like there are ids in the bigger structures, such as the immune cells recognize foreign cells/DNA, or the literal "id" tags in the DB, the logos etc.); for the more micro entities this hypothetical "IDs" could be stored in some, yet unknown, layer of existence and representation of the Universe, which is however accessible through the Will. The "causal IDs" may be relevant also for synchronization of the sensation of the wholeness, as there is lag between the parts of the system. The "wholeness", the unity may be experienced, felt, detected in some "out-of-time", more precisely "out-of-the-internal-universe-representation" time, where all the parts of the "causal ID" systems are sampled/parsed/connected at once. It might be just a speculative "fantasy" for the "real" physical Universe, but this is emulated by the brain/mind which is collecting all sensory inputs and its models/predictions of the world in all modalities and scales of time and space and is constantly building a coherent picture which is consistently merging and synchronizing all of them and keeping their continuity as much as the system can. One of the prices for the unification is the lag.

The "Causal ID" idea may sound as some kind of Dualism, but whether it is viewed as a division, thus whether the Will and the Representation are divided or united depends on the definition. The "Causal ID" could be another way to address the Schopenhauer's Will, the causal forces, which in more abstract and cognitive domains are addressed as reasons and all are felt by the entities, the agents, the minds; one important native property of the causal forces and reasons could be to communicate the unity of the systems, so the causal chains, records, dependencies, correlations, connections, networks could be some kind of "first-class citizens" as in the programming languages, they could be some sort of addressable "objects" which identify and wrap-up them in a whole, where this, as in other cases, could be

hierarchical, nested, overlapping.

If there are such causal-continuity identities, the "care" view of Michael Levin's "cognitive light cone"* could have objective backing. I agree that the more advanced/powerful causality-control units are ones which predict and cause at a wider spatio-temporal-modal-domain-…-… ranges and their development, and the development of general intelligence grows with these ranges (one of the range is to the sub-scale, micro: higher precision in the micro space, manipulating molecules, atoms, electrons). However the agent/mind itself "knows" or rather "it believes" or it may *express,* it may communicate, say, claim that it does, while actually its "real" goals could be different, or they could recognized, factorized etc. differently.

The "reality", the verity of its pursuit of these goals is questionable and is decided by an observer-evaluator.

M.L. illustrates it with the size of the cognitive light cone and "care" radius of a bacteria, a small insect, a dog, a human—different humans etc. , however this needs to be objectively measured which for abstract goals of humans, such as "saving the humanity" (especially if it's not defined from what), or "saving the planet" etc. is questionable, for several reasons:
***First,** the self as a mind is not unified*,*
***Second**—humans and other agents can pretend, can be confused, can lie etc. and the recognition/sampling, the system boundaries/nature/parts recognition and determination is tentative, it depends on the observer-evaluator (e.g. when one makes a donation for charity: does she do it for the good of a poor sick child from a third world country, or she does for indulging herself for her sins—guilt; or for her own pride; or because she believes that somebody expects that she should do it; or because she has spare coins; or … or all is valid at once or it is unknown: these concepts of guilt, pride, expectations etc. exist at some representations, in some minds, which are not objective and/or can't be reliably decided. Also, the further the goal post, spatio-temporally-abstractly, the more questionable is the direct causal relation of any action, and the more it depends on the orchestration of a bigger ensemble of other events, of the will of other, a bigger range or wills, goals etc. That's true either for specific goals, such as reaching somewhere in the physical space, or abstract; the wider the cognitive light cone, proportionally more other causal forces who impac the target goal and at best: the resolution of the causality-control and perception, with limited resources, would be lower.*
***Third—**this goals exist in one's mind, they are subjective; they get objective expression with particular actions, choices of the mind/agent—do this instead of this, because this is believed to be supporting the goal/state "A" instead of the others can be objective; the trajectory, evaluated in the future, which is matching the target one, approves objectively that the respective far-reaching goal is approached etc., however it is the agent who evaluates itself and/or another evaluator who does, that other evaluator again may be confused; related to the second point—while the more obviously measurable effects seem easier to compare, the more abstract goals often could be mapped to sets/sequences of more specific and less abstract and nobel ones, all complex goals could be sequences of simpler ones with shorter radius.*

- The lack of unity of self: see “Analysis of the meaning of a sentence…” from the TOUM (see below); also *“Nature or Nurture: Socialization, Social Pressure, Reinforcement Learning,* ***Reward Systems: Current Virtual Self—No Intrinsic Integral Self, but an Integral of Infinitesimal Local Selfs—Irrational Intentional Actions Are Impossible- Akrasia is Confused****—Hypothesis about Socialization and Eye-Contact as an Oxytocin Source”* http://artificial-mind.blogspot.com/2012/11/nature-or-nurture-socialization-social.html

When exactly the subjective “consciousness” or “Mind” appears?

Like other researchers argue, the "*Theory of Universe and Mind”* also discusses that the question when exactly the subjective “consciousness” or “Mind” appears can hardly receive a sharp objective answer (see “Man and Thinking Machine: Analysis of …”, 2001), also the brain goes through a rapid development in the childhood, the visual system preserves very high plasticity for the whole life etc., consequentially the concept of “realness”, as the concept and the perception of many other things, also constantly adjust.

Take images for example, video games and cinema. The eye of the players and the viewers gets more sensitive and the requirements to engage and attract them, as of fine detail,now are much higher than in the previous stages of the development of these technologies. Visual effects which were spectacular and were accepted as realistic 20 or 30 years ago now may look cartoonish. We can bring in the optical technologies and correcting glasses analogy here again, or to add also the cataract-removal experience, or cochlear implants being turned on for the first time; or the experience after a loss of some sensory or motor modality or its quality due to a trauma or stroke etc. We could extend this to the learning of concepts and broader shapes, relations, logic between the parts of the world.

Both the low level features of reality and realness, and the high level features and high level perception do change —the way we sample and choose the input, how we reflect about the world and the features; a viewer/observer/evaluator could understand it differently and pay attention to different phenomena, relations, scales etc. For example a child seeing “The Matrix” may see only fast action, flashy visual effects and hear the cool techno music and the noises, while when she grows up she may be interested in the friendship and love relationships and then in the philosophical implications of the pictures.

Years ago a friend of mine shared an observation that his parents didn’t care much when they switched from a CRT standard definition TV to an HD (1366x768)— they cared more about the semantics, the content. In the interdisciplinary work *“What a man needs? If you play by the rules you’ll lose like the fools.”, 2014*, near the end of the work Todor Arnaudov reflects on the resolution of the video images and the feeling of “realness”:

*** Какво му трябва на човек? Играеш ли по правилата ще загубиш играта!,** Тодор Арнаудов, 2014 сп. “Разумир”, бр. 1https://razumir.twenkid.com/kakvomu.html

(Google Translated + several corrections)

(…) The resolution of the view and the generalization
Where the advances in visual technology are leading to?

Video

Twenty or so years ago we watched movies in our homes on VHS tapes, re-recorded in abominable quality, probably 100 lines of resolution, 200 at best in later years(?); flickering, low contrast, etc., equivalent to only a few tens of thousands of pixels.

Today, screens showing 1080 lines, 2 million pixels, are popular, and 4K video is about to be introduced: 8 million pixels.

After watching any of these formats, however, the viewer's mind is not left with a significantly different memory.

It is only during viewing at the higher resolution that a greater sense of momentary reality is experienced (84)(269) and this may somewhat alter the path of attention in a random direction, which is not always desirable(240). However, our memory of reality lasts only a few tenths of a second, or even only as long as our eyes are open and we observe the image.

After that, all that remains in the mind is:

that there were explosions, car chases,
that the main character's car was a certain brand,
that someone played the main role,
that he had a romantic relationship with one of his heroines,
that the effects were cool
that we liked it
how we felt (fear, joy, pleasure…) etc.

These are the meaningful memories, the things that remain permanently in the human mind because of the way it works: in its more evolved part it summarizes, generalizes, works with small-volume compressed representations (103)(104)(203),—and still uses—his "old" parts (emotions) and connects the concepts with them.

The generalized meaningful concepts do not change with changing resolution, distinctness, of sensory input. A face is still a face, both as a miniature picture 60x80 pixels and as a huge poster of 20 million x 30 million. What is important and "meaningful" to a person in both cases are things like:

class: that it is exactly a face, not a horse, a car, a stone, a dog, a knife, a gun, a house, a lamp, …
do we know him—if so—who is in the picture?
gender (another class): male, female, transgender, transvestite
approximate age (another class): baby, child, adolescent, youth, … adult, 70-year-old, 90-year-old, …
emotion (…) : joy, sadness, fear, astonishment, surprise, disgust, …
does he have a moustache, beard or not
of what race, people, …
does he have a **hat**
does he wear **glasses**
is there a "**third**" **eyebrow** …

Etc.

The specific countless meaningless pixels that recorded all the pimples, moles, pores, cells—they don't matter—the details repeat, they don't matter for classification, and they can hardly be shared between people in any way other than through a picture that the mind can't usually output back. See (269)*.

The width of the information channel (203) of human mind [consciousness] is such that only generalized representations can be transmitted. [in a reasonable amount of time with the "built-in" effectors.

- Now even our mind can more easily output back its visual memories by reconstructing them via generative models, but still we input only a tiny amount of directions in the prompt.

Unless there are some direct "detectors", "connections" to reality, which suggest "a mind" that something was "real"— maybe there are such detectors at "visceral" level, in the texture of the living substance, the **realness** is recognized by **comparison** and **match, thus difference to a template that existed or was created previously.** The template could also be **chosen** previously.

The "simulation" called "Matrix" is a simulation only for the selected kind of system sampling and segmentation, from another one which doesn't know how "it should be" it is just the reality.

As of "simulating" or "modeling", there or elsewhere, it can be viewed as just another layer, "another level of dereferencing" as in the proverb about programming saying that *"All problems in computer science can be solved by another level of indirection".*

The same goes with "adding another layer" in the ANNs—for improving the quality in fine-tuned specialized models or for an overall higher performance,—or another layer in the ranges/spans/scales of matter: where different entities can "exist", i.e. be "observed", measured, sampled, evaluated, recognized, detected by a given other observers/evaluators that also "exist".

The "existential crisis" which some people have from some *"simulation arguments"** also may come from not taking into account that **all memories also turn into records as if they were not real—**not just the pale memories from the video images.

More **"Completely real"** or more "arguably" real sensory-experiential, "whatever it means", "living in the current working memory, default network …" or other "magic", is/could be/maybe … only the very-current "model" or "state"— however again whatever this "current model/state" means in "consciousness", "phenomenological experience", qualia, awareness—here again we face another choice of *"what we really mean"* by "*real*" and we stack many semi arbitrary choices, do we mean the same thing.

Some minds take some length of some “conscious processing”, be it “some” “integration”, be it some spatiotemporal slice of the operation of a brain*, be it 0.25–0.3 seconds between saccades of which Karl Friston mentions in his 3-rd appearance in MSLT; or it could be a longer period of several seconds for the “rehearsal” audio-speech memory (the phonological loop etc.)—the length could be different for different persons and in different states of alertness, tiredness and other conditions—so the “range of reality” shrinks or expands; the time span could be bigger, to include the short-term memory as well; or sometimes a shorter period for sound, the “sound tissue” of up to several hundredths of the second. The details, the elements, the “tissue” of the experience also differ—some observe, notice and reason about both tiny details and broad generalizations, others see only somewhere in between or see one part of the spectrum etc. Also, apparently the “realness” in different modalities is connected with different lowest level “tokens”* (transformer’s terms and any components, parts) and perhaps it is related to the presence of different lowest level nuclei/subcortical structures which first process them.

Back to reality or back to our mind or “they are both”?

Note, that the sensory integration of anything, of any input happens in a mind of an observer/evaluator, that is supposed to “exist” at some level, and the “hard problem” of consciousness seems to be unsolved, “existing” in a loop.

Doing anything while “not existing” seems a nonsense, not just the superfluous Descartes’ *“I think, therefore I exist”* : to be, or to do “anything” is enough, being measurable, detectable, means that *something* “exists”; however ***what exactly* exists,** what are its *actual* properties and do they match the expectations, how it/you “really” behaves etc.—how does it match to what— are different extensions of the questions and the problems.

For example let’s review a game world that depicts a city in a given country, either similar or different to the original—a big range could be drawn here as well, as even a text-based game could be “similar”, and texts/books are also “rendered”/simulated by the reader’s mind and they may reflect an actual world or a possible one, and the books could be filmed etc.

If the “real world” is described with text of particular syntax, semantic, lexical (vocabulary), logical “complexity” (measured in some way) it would match some “virtual” world which has a more detailed native/highest resolution representation and is then described at corresponding level of detail in text. The respective parameters would be indistinguishable.

What would be true about the textual story/textual game, would be that *“this is not the real city”*, but a “model” of it; yes, it is a *virtual city*, however that *doesn’t make it “not real”* as of “not existing”. “It”, “this world”, this instance of it, is located in the space of coordinates of having that game in this computer, running it, watching the screen/playing etc. that is its “realness”. It is defined by the trajectories of the properties or changes that have happened, in the actual physical states of the computer and the screen etc. which we assume that are “real” etc.

Even if this world is described in “non-dynamic” text*, in a book, that “world” also exists, first for example as a representation that has **a potential** to be “rendered”, and then as

run/simulated world in the mind of the reader who makes it dynamic by reading it—the process of reading is animating, moving the text literaly, by "sliding" the text as a tape, like the Turing machine,— then the mind interprets it in the imagination and fills-in additional potential/possible "data points", physical laws and relations. **The text and anything perceived and thought of by a mind does exist with the existence of the mind of the observer-herself*.**

- ***"Non-dynamic" text* — the dynamics is conditional,** it requires anchor definitions and evaluators; **the video game worlds also have static representations** when they are stored before being run or if we observer a "frozen" frame; their space also could be represented as static, as a network of all possible game states and one is just sampling particular coordinates, points, vertices, "rooms" (imagine a text-based quest) etc. within that representation that can be expanded as "static" from the start. If one chooses to account for the state of the *screen* or the state of the *RAM* etc. and to think of them as of "existing", "actually existing", "living" only regarding the content that is projected on the screen in a given moment; the representation which is in RAM; in the focus/attention-span of the observer—for a given spatio-temporal range—etc. (Again, for all: with a given choice, anchors, borders, resolution etc.), then the game could be seen as more "dynamic", with a focus window scanning the different portions of the potentially expanded/unfolded representation. However the same could be imagined for the books, see above.
- ***Existence with the mind of the observer*—**that reminds me the intro from a video lecture in MLST from a lecture, where the lecturer ridicules panpsychism with suggesting that "perhaps the rocks are conscious?" etc. Well, while you think about the rock, **"it"** does exist as representation in your mind, your mind is activated in a way that is affected by that concept, memory etc., so if you do have that mighty precious special consciousness* ability, you now play rock and **"it" has gotten your consciousness. You are a rock now!** (* The "consciousness" often seems or is treated like yet another **ape's** social ranking logo/"diploma" and a way to segregate higher ranked ones from the lower ranked one. See this discovered even in the *Theory of Universe and Mind, 2001–2004*)

Axioms, lowest level representations/instructions, the primary/most basic rules etc. are artifacts of any form of (computational) irreducibility* for a given system, which is not a problem for solving problems and for extending what these systems can do by developing longer are more sophisticated compositional programs which employ the lowest level instructions/laws/code/representations etc. In order to exist, the lowest level has to be mapped, to communicate with some properties, parameters, states of the "lower" level universe, from its POV, it's based on some "lower physical" property, such as the bits in RAM or the CPU being electrical potentials of particular "capacitors" or the state of a "flip-flop" circuit, which are particular areas of the surface of a chip, which if electrically measured by a particular procedure would indicate particular voltage or particular current would flow in particular parts of the electrical current; and a "capacitor" or a "flip-flop" on these chips is a "transistor" or a circuit of several transistors, which are particular … etc.

The kind of irreducibility I refer to here is different from the one requiring to run a simulation in order to produce particular outcomes (see S.Wolfram). This one is about the elements, the existence of some "bottom level" of segmentation and representation in a given system.

Another reason why Goedel incompleteness is irrelevant, as well **"the halting problem",** is that the systems are hierarchical, multilevel, multiscale, multirange, multidomain and the (system, computer, representation) at the lowest level from the POV of some currently observed/evaluated one is **supposed to never crash**, as long as the current level exists. The lower levels are supposed to "catch" and deal with any possible **"exception"* or an "interrupt".** Note that the current level is even **undefinable** for an observer if the lower one doesn't exist, and the "axioms" at the lowest level are "just data", representation that do not need any further prove or the prove is direct—applied in one step, one check,— or there's no one, that's "the bottom".

Note that this **maps** to a debate of some "existential crisis" people who are worried about the fact that the transformers are just **"mighty hash tables", i.e. maps.** The final step of processing is direct, obvious, apodictic: a function, a mapping and this is not "wrong". Not that the cognitive models do not need depth, hierarchical decomposability, extension, connection, "grounding" to other concepts, nodes etc.

IMO the "halting problem" also depicts that infinity is not an empirical concept, as if I'm not mistaken Pedro Domingos? points out in MLST #096. (Some people may consider "**empirical**" as a proof for "being real", but what is "legally" considered "empirical" should be defined, and IMO that's an overgeneralization, either because sensory input is also prone to illusions and errors, it may map to something different than what was interpreted, and because **there are non-empirical concepts**, where "non-empirical" is such in the sense of the "conventional" explicit sensory matrices; even without such explicit sensory matrices, there could be sensations, feelings or non-empirical concepts which are still "experienced", cognized as well, and we can think of them—i.e. including the "internal sensations", the "qualia/consciousness/the Will"; abstract concepts at different level of "pureness" which may have mappings or their instances can be rendered as representations and applications in the "normal" sensory-matrix-addresses-based sensory world, but their essence could be procedural, a sequence of transformations, directions etc. I think Kant in the "Critique of Pure Reason" gives an example of these abstract concepts with the "triangle" and that the abstract concept itself is not of any single triangle, but of the space of the possible ones.)
See also "*Embodiment is just coordinate spaces, interactivity and modalities—not a mystery"*, T. Arnaudov 2011: https://artificial-mind.blogspot.com/2011/12/embodiment-is-just-coordinate-spaces.html
[from computational, not subjective POV]

* The "Exceptions" are discussed in the Theory of Universe and Mind, e.g. "Concept about the Universal Predetermination 3", 2003 and in the specific article *"Abstract theory of the exceptions of the rules in computing machines. A Theory about the Control units and the Control"*, 3.2004:* https://web.archive.org/web/20041020165359/https://bgit.net/?id=65835

Today (5.5.2023) I discovered a “lower-level-of-abstraction?-physical-systems-theory” work which seems related, without the digital universe POV, and it also discusses this phenomenon of lower levels dealing with the dissolvement of the higher level representations—in their context/vocabulary: level of development and organization. In ToUM there is also a chain/degree of development of the Universe, embodied in causality-control units which get better in prediction, achieve higher levels of generalization/compositionality etc. and aim at becoming closed system, i.e. to be able to predict their future exactly with probability one, to become independent from the rest of the Universe (something they seem to be unable to do in the strict/”real” sense, unless, hypothetically, they are or become the whole Universe). Note that Georgiev brothers? define and discuss the properties of open and closed systems, in ToUM a “closed system” was defined as one that doesn’t need external information and reach to maximum possible prediction, for example the computers in ToUM POV are closed systems when they have the entire input data, the execution of the processes, the instruction flows, is prediction with probability one. Recalling the LLM/transformers, note also that the merits of their technology is diminished to *“just predicting the next token”,* however computers in general do the same too*: **computation is prediction** of the result of the current instruction, their tokens are their **instructions** and the next token or “next word” is the output of the current instruction and the next address; **causality** can be seen the same way: *“the Universe just computes the next state—the next token, one at a time at the highest resolution”** and all are the same and they all can be represented as **mappings**, **functions**.

- At least that’s the assumption I make, and that’s our “objective” perception/POV of our mind as sequential observer-evaluator of the states/the sensory-motor inputs-outputs.

On the other hand, the living organisms, cells are known open systems and the claim that higher levels of causality-control units (higher degrees of development of matter, universe, …) aim at becoming more closed may appear counterintuitive; similarly and paradoxically, in this regard, one of the insights which suggested me that Universe was likely deterministic / digital were the laws of the random numbers—in order the random numbers to follow the proper distribution, I assumed that there must be a “**controller**”, a “master observer” that watches all events and balances them, otherwise there would be anomalies, for example one could throw 50 or 100 or 1000 consecutive “head” or “tails” of a coin; in reality one could hardly throw even 10. Sure, some mathematician may suggest that on each throw the probabilities are equal etc., however these are theoretical claims and even then—how the odds are “initialized” before and after each throw, in order that to be true for *all* possible coin-throwers and random events in the whole Universe, and why this is true at all? There must some common “initializer”/”watcher”/”balancer” to take care of the proper distribution. Bobby Azarian mentions a similar event-correlation reasoning about the tendency of the Universe to what he calls “Omega point” in the podcast
https://www.youtube.com/watch?v=_s-ziTJ3KzI

* G. Georgiev, and I. Georgiev, The least action and the metric of an organized system. Open Syst Inf Dyn 9 (2002) 371–380. (cited in Michael Levin's: Technological Approach to Mind Everywhere (TAME): https://www.frontiersin.org/articles/10.3389/fnsys.2022.768201/full

"This state is called organization. If the action is increased by changes in this organization, the system is destroyed, or **drops in its level of development** as measured by the action. This is a first of a series of papers that will have as a purpose the description of the process of complexification."*

Final Notes

1. AFAIK Schmidhuber also believes in the Universe Computer and that the randomness is not really random, but he calls his general intelligence model "Goedel Machine".
2. Bobby Azarian's theory in *"The Romance of Reality: How the Universe Organizes Itself to Create Life, Consciousness, and Cosmic Complexity", 6.2022* (as of the intro and the podcast "*The Romance of Reality | A Unifying Theory of Everything", 29.4.2023,*
   The Truth w/ Carlos Farias https://www.youtube.com/watch?v=_s-ziTJ3KzI) also sounds analogical and related to the ideas and conclusions of "Theory of Universe and Mind" (2001–2004).

3. In the discussion in MLST community there's an argument that some combinations of words would be unseen and couldn't be computed; to me this is irrelevant, just because one doesn't have to predict/compute everything using that same method which is used now and has these weaknesses, e.g. the "tokens" as they are or just one set of tokens and one unified dataset. If you do not have appropriate prior data /experience /model you just have to use another model, a more general one, to reduce the resolution of causation/control or perception (or to increase it and to predict other representations), to switch the Context {K}, that's a concept from my unpublished works from the early 2010s—more on it later—and view the problem/future from a different POV where you do have models etc. In general the ultimate, most powerful models of the universe and mind are supposed to be causal, not (just) probabilistic, probability, i.e. uncertainty, lower than P=1, is when the model is yet "young", incomplete (see a discussion in the same Discord channel from mid April, Todor's (Tosh) points about creativity and P() = 1) or/respectively if the prediction horizon is too far away for the capabilities of the model for the moment, thus the prediction resolution is too low/it's below complete causal model*—there is not enough experience or reliable data and/or appropriate causal forces which are at the command of the agent/mind ("the best way to predict the future is to invent it", Alan Kay—one of the ways t predict the future is namely to act it, to "force" it to happen; in Karl Friston's theory that's the "Active inference")

- Some people criticize the "mechanistic" view to the Universe and may make that objection for the above reasoning about causality: "the Universe is not mechanistic". I don't think so either, to me it is **"informational", computational, "informatical".** The informational/computational systems = transformational and when they are complete, they include all levels of the virtual universe/CCUs. These systems don't

have to obey the simplistic "hard frames" applied to the typical "mechanistic"or "mechanical" phenomena, because "the back end" of the Universe Computer can write or read anywhere, transparently for the evaluator-observers inside the universe. Action at a distance is not spooky[48] as well, when the Universe is a computer.

4. Yes, "Ada" has inspirations, influence and references to "The Matrix", as well as to other pieces, which also has reasons in the story. I recently rewatched it, the first part, and I noted the Elevator scene. There was a caption about "Service" for a moment. I don't remember that I have had realized that when I wrote the novel, but it seems I've remembered this tiny detail.

5. It's interesting that the authors, maybe brothers, are also Bulgarian, I didn't know their work before writing this article. "TAME" and other concepts regarding "Scale-free cognition" and "computational definitions of self" of Levin are also similar/analogical to the Theory of Universe and Mind's concepts—the "self" there is causality-control unit, there are "subunits", prediction-control, different ranges/resolution at different scales of the development of matter and the systems (his "cognitive lightcone") etc.

6. A github repo dedicated to the ***"Theory of Universe and Mind":***
https://github.com/Twenkid/Theory-of-Universe-and-Mind/

7. Regarding the "Causality-control units" and "subunits" in ToUM, compare terminology to M.Levin's "subunits" in the TAME paper.

8. **Steven Wolfram also refers to "constructive mathematics",** constructivism when discussing his image of the Universe as a cellular automaton which is rendering the states. His definition of understanding expressed in recent videos (e.g. MLST, Lex Fridman's podcast, April-May 2023) agrees with mine in "Man and Thinking Machine", 2001 and later, e.g. "Universe and Mind 3", 2003. Maybe this is a general computationalist and patternist view: to make the input computational, to split/segment it to parts which can be analyzed, processed etc. and to do whatever is possible to be done with them etc.

9. **The realness of the "persons" from the generated images.** There is a popular website called **"Thispersondoesnotexist",** demonstrating the capabilities of the StyleGAN generative model. We also read about generated images described as "*this is not a real person"* or *"not a real scene" etc.* These are the same confusions about **realness**: in any case, even if they were photographs, the photos are only "images of…",they are not "the person" that is on the picture, either if it "does exist" (is photographed directly) or it's synthesized by a generalized model of an image of a human face. Also, of course it is possible and highly likely that humans with looks similar to the generated images **do exist**, that's why we perceive the images as realistic and if one searches and if decreases the resolution/strictness of the match, the features, she probably would find many matches with "persons that do exist".

[48] "Spooky action at distance" – the quantum entanglement. The explanaition in TUM is given in UnM4, in the slides of the lecture on TUM at the AGI university course, in "Stack theory is yet another…", 2025.

- **Add about Anthropomorphizing … … note 8.6.2023—cite from the article.**

10. While I agree with many points and reasoning of Arthur Schopenhauer, I was not inspired by his philosophy in the 2001–2004 period, as I studied his work a decade later: I first read the "Aphorisms …." in late 2004 which are not so deeply epistemological, but I noticed that his way of thinking and style of writing were similar to mine and it was easy for me to understand it. Maybe in late 2010 I found in a bookstore 3 out of 4 volumes from his "The World as Will and Idea" (the first was missing) and bought them all, however I didn't read them at the time, I only glanced some pages. I guess this article was written about that time: https://artificial-mind.blogspot.com/2010/10/arthur-schopenhauer-world-as-will-and.html
I mention the matches to my theory which I have noticed, referred there as "Teenage theory of Mind and Universe". The time for studying Schopenhauer's body of work was about 2014, when I read all that I could find in English and Bulgarian, part of the latter in two different translations: thousands of pages in total.

11. **There could be more fine-grained "modes" of "rewards/experience" than just "physical/cognitive".** It seems possible that the "physical ones" could have separate dimensions for different basic needs or there could be several subdimensions. Arthur Schopenhauer argues that the common between the emotions, the "irrational" is that they can't be generalized. Also, the neural substrate required for registration of the emotions by mammals require the proper operation of particular parts of the brain, where the "higher" parts have to map/interface/connect their operation and representation to the lower one, they are system, a whole.

I've argued that the "higher" modules of the brain, under the Triune brain nomenclature, reiterate the same functions of the lower ones, there was support evidence regarding the Anterior Cingulate having similar patterns of activation to the Prefrontal cortex when facing uncertain situation, which requires more planning, reasoning.

However the "lower" brains have less cognitive and representational capacities and simpler kinds of memories. While the cognitive, or *the most cognitive* brain/mind/stage of processing, connected with a requirement of a functioning neocortex, can recognize "unreal" inputs by comparing them to some baseline, the lower modules, lower resolution ones might be unable to "understand" what "unreal" means at all, because they just "feel" it directly; possibly with a difference, the function governed by the Cingulate (assuming the rest of the brain is in tact as well) may have some better sense of "unrealness" compared to the lower subcortical structures, as it's more developed and more closely connected to the neocortex, and behaviorally the Cingulate is linked with the recognition of errors.

Also, there could be different experiences of/for "reality" involving these different "modules" and/or "stages": are they unified, are they switching, is there an "attention" which is switching between them or distributing their weight?

Logically, the most basic form of "unrealness" could be the detection of the mismatch between expected/wanted/predicted and the measured/sensed with at least the two different kinds mentioned earlier, and possibly in different modalities and different aspects and brain structures and systems. Note the epithets "incredible", "unreal" etc.

…

TOUM argues that the more highly developed, the more advanced Causality-Control Units have more capacities for indirect interpretation of the input as just data and ignoring irrelevant/undesirable and interpreting it “as they wish”, while the less advanced, less developed, “lower” ones (lower levels in a cognitive-control-causality-prediction hierarchy/system) interpret the input, the data, the sensations, more like direct instructions and maps to actions.

In bodily terms that’s reasonable as if the lowest levels for the body/the physical “implementation” are the most inflexible, they are like the “hardware” and errors in their operation more quickly lead to damage and death, including of all the higher levels. Recall from “Concept about the Universal Predetermination”, 2003 and “Abstract Theory of the Exception of the Rules”, 2004 that when an unpredicted, undefined “instructions”, states are reached in the virtual universe/causality-control unit, a lower level one takes the control, as by definition the upper level is defined by instructions from the lower and the lower one should have a more complete coverage of the possibilities for action and the “physics”. The lower level system takes care for the errors to find a work around and correct the issue, generate another correct code, repair broken tissue, compensate some chemical disbalance by altering other parameters of the system etc. However if the error is at the lowest level of the system/virtual universe/causality-control unit/agent, it irrecoverably dies.

On the other size, this is unlike the higher level causality-control units which are increasingly virtual, imaginary and flexible and errors/mispredictions/”impossible” states in their existence and “experience” are just captured as “exceptions” at a lower level Virtual universe and resolved, the higher level is corrected and respawned etc. See TOUM and in particular “Abstract Theory of the Exception of the rules in Computers. A theory about the control units and the control”, 2004

…

12. Triune brain theory is challenged, but it doesn’t matter

There are objections against the triune theory; while it is true that the segmentation of a system is by default dependent on the observer/evaluator, the criteria, the choice of the components and the segmentation and which are the system elements, where are the boundaries etc. and that the “higher” functions of the cortex are impossible without the subcortical and the brain stem— these functions are also impossible without the whole body or a respective “surrogate” and eventually: **without the whole Universe: without the Universe nothing exists!** etc.—therefore it could be argued also that “there’s no brain”, because brain is just a part of the body, and “there’s no body”, because it’s just a part of the Unvierse.

The selection of the components, the system recognition, the logical/system-level division of these parts and areas is based on anatomy, some choice of “what is different enough”, what is “connected enough”, either physically and/or functionally, brain cortical areas and nuclei “firing together”, having “projections” (neural pathways) etc. —in machine learning and informatics this segmentation, analysis and synthesis process, is a kind of Clustering—; there is a choice of particular methods for analysis related to functions, which repertoires of behaviors and capacities get deficient if the respective parts are damaged or

lacking etc.*

However when working together these “modules” are a whole, their “independence” gets evident when some of the stages is damaged or suppressed. For example just high level of stress and panic may lead to “temporal stupidity”, starvation of the neocortex, less blood supply—human behavior gets less “rational”: nervous and panicked people make more mistakes, speak worse, can’t concentrate, get very aggressive etc. However the cortex is still there, it’s suppressed and working at a lower resolution etc., but it is not removed; if it was completely shut off, the stressed subjects couldn’t utter a word. A longer term chronic cortisol exposure causes shrinking of the hippocampus etc. Alcohol intoxication or frontal dementia/atrophy of the Prefrontal cortex (PFC) may have similar effects to the capabilities for abstract thinking, executive functions, as the PFC integrates and can inhibit the rest of the brain, and may make the patient exhibiting a more aggressive or just “socially inappropriate” behaviors and less “self-control” etc.

13. The experimental **“atlernative formatting”** is not a notation of the mentioned language of thought “Zrim”, it was just an improvisation on the fly due to the complex sentences.

14. The popular dismissal of brute forcing which leads to apparently intelligent behavior as not-intelligent doesn’t pass a deeper analysis (11.9.2023)

An answer to the current MSLT episode with M.Mitchel; the story about chess, in the early days the AI-field believed that if it is solved it will require human level intelligence etc., but it was discovered, that brute-force search works, so “it’s not intelligent”.

Similarly with the transformers, current mainstream neural networks trained with backpropagation etc.

According to the mainstream though everything is a result of the “evolutionary process”, which, according to the yet official position, is a “brute forcing”, it’s a search, exploration, testing etc. (However when there’s DNA already, that search is already in a narrowed space, the individuals can meet and mate with other already selected individuals etc., it’s “digital”, discrete, quantified)

Also, as suggested in “Man and Thinking Machine”, 2001, that what’s often attributed as “human intelligence” is actually the intelligence of the whole humanity, which in the end is the whole Universe—it is not “the brain” or an individual, because the brain can’t exist and do anything without the rest of the body, and that body needs environment and finally: a whole universe. M.Levin: ~ “all intelligence is collective intelligence”, also F. Chollet? etc. understand that there are “priors” and that intelligence is connected and dependent on the environment, the context, it is “ecological”.

Even just because of the above arguments the human intelligence can also be seen as, or it *is* also “brute forcing”. It’s actual material implementation requires the whole universe to run—as anything else—and where the boundary, the limit of the sampling is set depends on the observer-evaluator.

The human intelligence can be seen as brute forcing also if it’s considered “intelligent”, as still a huge amount of sampling and processing is required even in some abstract purely computational representation.

The cerebellum alone is said to embrace more than 100 billion neurons dedicated for

computing and synchronizing the physical motions of the body and the body first explores these spaces, similarly to the RL/genetic learning bots in simulations. The same goes for all cognitive activities in their respective representational spaces.

…

In all cases the "bruteness" is a matter of degree, and also in order to *implement* it, equivalently fine and deep technologies are required. So computers "just compute", LLMs/NNs are "just matrices". Well: how the machineries for that are built? How fine are their transistors, how big is the code base of the operating systems, who huge are the dataset.

Why this is not examined and included in the calculations of the complexity?

The observers-evaluators *choose something which fits their aim to diminish* an object, therefore this is "it". "It" is their selected simplified representation.

Is it "intelligent".

(According to Theory of Universe and Mind and others, such as … "Omega point"?, FEP, "Diverse intelligence", Cognitive lightcone, …—there is some form of agency in all levels of scale)

By the way this is also the **"it, but what is it?"** problem (the term is made up now). "It" is related to the phenomenon which was investigated in "Issues with like-dislike voting…"—the viewer liked *"that",* but she doesn't specify *what is that;* sometimes it is known, sometimes it is not.

References

Other References and Links

- https://en.wikipedia.org/wiki/G%C3%B6del%27s_incompleteness_theorems
  https://en.wikipedia.org/wiki/Peano_axioms#Peano_arithmetic_as_first-order_theory
- https://www.mdpi.com/1099-4300/24/5/710
  Biology, Buddhism, and AI: Care as the Driver of Intelligence
  by Thomas Doctor 1,2,Olaf Witkowski 2,3,4,5,Elizaveta Solomonova 2,6,Bill Duane 1,2,7 andMichael Levin 8,9,*ORCID
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6923654/
  The Computational Boundary of a "Self": Developmental Bioelectricity Drives Multicellularity and Scale-Free Cognition
  Michael Levin1,2,*†
- Architecture of the brain of the mammals—from the AGI course 2010, 2011, see slides 23, 24:
  https://research.twenkid.com/agi/2010/Brain_Architecture_22_4_2010.pdf
  Basic emotions (modes of operation): *Seeking, Rage, Fear, Panic, Play, Lust, Care* and their respective associated systems in the brain (that could be partially outdated as it is based on papers from 2005,2006 papers)
- Supports for the Hypothesis of Higher Brains Re-iterating the Functions of the Lower Ones—[that] implies that the function of the higher ones could be induced by understanding the functions of the lower ones and extending them
  http://artificial-mind.blogspot.com/2013/08/supports-of-hypothesis-of-higher-brains.html
- Proposal for Directed/Guided Evolution of a Cognitive Module/Algorithm by Step-by-Step Modification of a Basic One—from Archicortex to Neocortex
  http://artificial-mind.blogspot.com/2011/02/proposal-for-directedguided-evolution.html
  See also Jeff Hawkins "On Intelligence", 2004
- Differences between Decorticate, Midbrain and Decerebrate Animals by SB Deshpande, https://youtu.be/8rI1XjII4og
- See also "The Neurobiological Underpinnings of Psychoanalytic Theory and Therapy", Mark Solms, 2018,
  https://www.frontiersin.org/articles/10.3389/fnbeh.2018.00294/full
  See also the comment about this paper in the github page of ToUM.
- "Ada" in sf.ludost.net (Bulgarian):
  https://web.archive.org/web/20041014172228/http://sf.ludost.net/autors/h/Todor_Arnaudov/Ada.htm (Chapter 2. is missing here)
  https://web.archive.org/web/20040611152740/http://sf.ludost.net/
  https://www.oocities.org/todprog/ada/ada.htm
- *Why anything should exist at all? What's the meaning of life/universe/existence…. ?*
  In the 2002 work from TOUM *"Letters between the 18-year old T.A. and the [43-*

years old] A.G.", with one of its main titles *"Concept about the Universal Predetermination"* once I argue/discover, that there is no need for having a purpose/meaning/sense.
* This topic is about the "existence", thus sort of "existential", however it is not "depressive", as it's not/should not strictly be experienced emotionally, affectively, as some "extreme" emotion such as despair etc., but it is felt logically and cognitively. "It doesn't matter that it doesn't have a meaning/purpose", this discovery may not change one's motivation, "will to live" etc. IMO ahealthy body, being, entity, mind, causality-control unit, which is not suffering and considering that it can not escape that state, is supposed to have "pleasure" in just existing, to have a fair goal in the existence per se. It doesn't have to have any additional special imposed enforced purpose, given by a personified master; to serve somebody, to "be useful for society" etc. It is even more extreme as A.Schopenhauer comments that even very old and powerless persons, who are suffering for a long time with many aches and disabilities due to their age and dissolving body, are still motivated and struggling to live more, besides the pain, to live on another day, another year.

- The cognitive science concept of the phonological loop : https://www.sciencedirect.com/topics/psychology/phonological-loop#:~:text=The%20phonological%20loop%20comprises%20a,representations%20in%20the%20phonological%20store.
- *mindcell—a building block of a mind; also a possible English term for the Zrim concept of{K}/Контекст/Клетка (Context, Cell); in this context a synonym of causality-control unit.

Resources for the original writings of Theory of Universe and Mind, the AGI courses etc..

@Vsy: Extract from the text as well.

1. T.Arnaudov, Theory of Universe and Mind's Github page, 2023- https://github.com/Twenkid/Theory-of-Universe-and-Mind
2. T.Arnaudov, The Prophets of the Thinking Machines: Artificial General Intelligence and Human Development*: History, Theory and Pioneers; Past, Present and Future", 2025
3. Т.Арнаудов, „Вселена и Разум", Теория на разума и Вселената – списък със статии от историческата страница на дружество „Разум" 2001-2004:
Original home page of "Society Razum" of *The Sacred Computer*:
* https://eim.twenkid.com/old/razum/index.htm * http://oocities.com/eimworld/razum и др.
Archived page from January 2005:
https://web.archive.org/web/20050112110951/http://eim.hit.bg/razum/
4. Т.Арнаудов, Истината, 2002; 2003
* The Truth, T.Arnaudow, 2002/2003
* https://chitanka.info/text/865-istinata
* https://www.oocities.org/eimworld/eim19/istinata.htm

5. Т.Арнаудов, Ада, 2.2004 https://www.oocities.org/todprog/ada/17-23.htm https://www.oocities.org/todprog/ada/
* The edition in an e-library (see chapter two in the comments section, it was missed in the upload) https://chitanka.info/text/866-ada
6. T.Arnaudov, Angel Grancharov, Letters between the 18-years old Todor Arnaudov. and the philosopher Angel Grancharov, 2002 (Писма между 18-годишния Тодор Арнаудов и философа Ангел Гръндчаров, 9.2002)
7. …
Bgit.net links – see at archive.org
* https://web.archive.org/web/20050112211003/http://bgit.net/?id=65395
* https://web.archive.org/web/20050114204244/http://bgit.net/?id=65835

* Т.Арнаудов, „Емил и Кибертрон“, сп. „Свещеният сметач“, бр. 20, 23.2.2003, – Обявяване на умотворско дружество „Разум“.
* T.Arnaudov, **“Emil and Kibertron”,** The Sacred Computer #20, 23.2.2003 – Announcing the “mind-creating” society “Razum” (Reason, Mind) – the invitation for partners, the first AGI start-up pitch, before the AI strategy 3 months later.
* https://eim.twenkid.com/old/eim20/emil.htm
* https://www.oocities.org/eimworld/eim20/emil.htm

Other materials from the world’s first university course in AGI (2010, 2011)

* https://research.twenkid.com/agi/2010/en/
* https://research.twenkid.com/agi/2010/
*https://research.twenkid.com/agi/2010/Todor_Arnaudov_Theory_of_Hierarchical_Universal_Simulators_of_universes_Eng_MTR_3.pdf (English) – Lecture slides for the AGI course, already the
*https://research.twenkid.com/agi/2010/Todor_Arnaudov_Theory_of_Hierarchical_Universal_Simulators_of_universes_MTR.pdf (Bulgarian – a shorter version)

Translations of Universe and Mind 3 and 4 from 2011[49]

https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Universe_and_Mind_3.pdf
https://research.twenkid.com/agi/2010/en/Todor_Arnaudov_Theory_of_Universe_and_Mind_4.pdf

* https://www.oocities.org/eimworld/eimworld13/izint_13.html - „Човекът и мислещата машина: (Анализ на възможността да се създаде мислеща машина и някои недостатъци на човека и органичната материя пред нея)“ = **Man and Thinking Machine: Analysis of the possibility of creating a Thinking Machine and some disadvantages of man and**

[49] They needed corrections of typos etc.

organic matter in comparison to it[50], 2001 (Български)
* Следващото еволюционно стъпало, https://eim.twenkid.com/old/eim17/sledvashtoto.htm 3.2002; 7.2002
* Схващане за всеобщата предопределеност, Мисли за съдбата, за разумното и духовното съзнание, и Мислещия сметач, 7.2002, https://eim.twenkid.com/old/eim17/predopredelenost.htm
* https://eim.twenkid.com/old/eim18/predopredelenost2.htm Писма между 18-годишния Тодор Арнаудов и философа Ангел Грънчаров, 9.2002… Вселената сметач, Вселена и Разум 2 (Схващане за всеобщата предопределеност 2)
* https://www.oocities.org/eimworld/eim18/predopredelenost2.htm …
* https://www.oocities.org/eimworld/3/25/pred-3.htm – Вселена и Разум 3, 2003
* https://www.oocities.org/eimworld/4/29/pred4.htm – Вселена и Разум 4, 2004

* **Вселена и Разум 5,** сп. „Свещеният сметач“, 12.2004 https://web.archive.org/web/20050124070240/eim.hit.bg/vr5.htm – Universe and Mind 5
* „Анализ на смисъла на изречение въз основа на базата знания на действаща мислеща машина. Мисли за смисъла и изкуствената мисъл “, Т.Арнаудов 2004 https://web.archive.org/web/20040402125725/http://bgit.net/?id=65395 http://artificial-mind.blogspot.com/2008/02/2004.html (some remarks)
* Абстрактна теория за "изключенията от правилата" в изчислителните машини. Теория за управляващите устройства и управлението., 18.3.2004 https://web.archive.org/web/20041020165359/https://bgit.net/?id=65835
* Abstract theory for the “exception of the rules” in computer systems. Theory of the causality-control units and causality-control, Todor Arnaudov, 18.3.2004
* “Analysis of the meaning of a sentence, based on the knowledge base of an operational thinking machine. Reflections about the meaning and the Artificial Intelligence”, Todor Arnaudov, 18.3.2004 (in Bulgarian; translated in English in 1/2010: https://artificial-mind.blogspot.com/2010/01/semantic-analysis-of-sentence.html - the paper is split in 4 parts, titled:
* Part 1: Semantic analysis of a sentence. Reflections about the meaning of the meaning and the Artificial Intelligence http://artificial-mind.blogspot.com/2010/01/semantic-analysis-of-sentence.html
* Part 2: Causes and reasons for human actions. Searching for causes. Whether higher or lower levels control. Control Units. Reinforcement learning. http://artificial-mind.blogspot.com/2010/02/causes-and-reasons-for-any-particular.html
* Part 3: Motivation is dependent on local and specific stimuli, not general ones. Pleasure and displeasure as goal-state indicators. Reinforcement learning. http://artificial-mind.blogspot.com/2010/02/motivation-is-dependent-on-local-and.html
* Part 4 : Intelligence: search for the biggest cumulative reward for a given period ahead, based on given model of the rewards. Reinforcement learning. http://artificial-mind.blogspot.com/2010/02/intelligence-search-for-biggest.html

[50] I have translated it also as “…Analysis of the Possibility a Thinking Machine to be Created and Some Disadvantages of Man and Organic Matter in Comparison” but then corrected it.

[See the multiple conceptual matches with the 2024 theory by Thorisson & Thalevi]
(…)

@Vsy: **To be continued and refined**
@Вси: Събери препратки и ги добави.

Appendices

Is the emergence emergence and emergent?

Appendix to “Universe and Mind 6” (UnM6)

Todor Arnaudov, ~1/2024[51] (inspired by discussions by M.Levin etc. – talks about experiment with evolving sorting algorithms of 6 lines of code)

How emergence is induced, perceived? States, results, data coming “unexpectedly” etc.

*** Comparing different domains, views, scales which are incompatible**. The classical example with fractals (in the video: *sorting algorithm*): the saying that *“how come so simple rules generate so complex phenomena”* etc.

1. The rules are simple only in one format of the representation *for a mind* with a particular sophistication in an abstract space – one complexity is traded-off for another.
2. The *actual* rules for the universe require *complete* definition of the implementation.
3. Even in abstract spaces, the fractal is unfolded in *memory* and in *time* and these operations add up to the complexity, if they are considered as well.
4. The *surprise* from that apparently *new* complexity (unpredictability or number of elements given particular segmentation/analysis, tokenization model etc.) comes from the evaluator-observer's inability to predict/see it, without that *additional* memory and computation (Levin admits that a science is needed for that), with some desired/expected precision etc. – the *unpredictability* is part of the "computational irreducibility" of which Wolfram talks; thus complete science, which could predict all details at the highest resolution of the output, may be impossible; in TOUM: the higher level/more abstract causality-control units predict with a lower resolution than the lower levels. The simple definitions are already compressed, they lose the detail of the originals. *However* as always this depends on the observer’s capabilities. What one system is incapable to predict or model “simply”, is obvious and accessable like a hash-map for another one.

* If we evaluate even the most complex program, executed on a single CPU core, as a linear algorithm, as designed by the programmer, at each or a given *“shortest”* possible moment, it would be as complex as a single instruction or the sequence or context with a given length (several instructions, a subroutine etc.) and the state of the registers of the processor, the memory, the I/O devices etc.; therefore *locally*, *momentary* and given a selected, chosen *range* and *resolution,* it is not more complex than *the simplest program (or maybe “any program”)* of length 1 or N instructions etc., or the difference is only directly reflecting the different types of possible instructions and the clocks they require etc. (MOV, CMP, JMP, DIV, …)[52]; the complexity is expanded and rendered through traversing all states and accumulating the differences. It is the same with the apparently simple algorithms for the observer. For a program, a sequence, a structure of any length or size, the observer-evaluator may know that the more complex program is more complex only if she can *remember the history/the whole*

[51] Written ~ 1.2024, a few short clarifications and edits on 21.9.2025

[52] A detailed review of this range would be complex. [21.9.2025]

range and evaluate it altogether. Otherwise, all programs are at most as complex as the maximum of the complexity per *"focus span"*, current range of evaluation etc.
* For a sample processor in such theoretical comparisons, it's more convenient and more explicit if simpler 8-bit or 16-bit CPUs are evaluated, in order to avoid or reduce the "spilling" of the execution stages and other complications etc.
5. The evaluator that estimates a higher complexity has a corresponding intelligence and a sufficiently large receptive/processing span and depth, which encompasses all range of details and finds "unexpected" or "beautiful" correlations (matches). Thus part of the complexity evaluation and appreciation comes from the mind of the evaluator and the complexity encoded there[53]. As in art, music, literature, science – one may have to be apt in the domain in order to understand and appreciate the technical mastery with a knowledge of the techniques (one can be amazed by being unable to imagine, predict etc.).
6. → **Local unpredictability, global predictability ...**
As explained in the classical TOUM, as early as 2002, which was first derived from the random numbers – the laws of the random numbers suggest determinism, or as some call it "superdeterminism" – in order the probability distributions of the averages to work, all has to be synchronized and all events to counteract each other. C.Fields, and M.Levin in 2022: " No finite agent can discover all of the causal influences that determine its own behavior, as global determinism logically disallows local determinism. Hence, global determinism assures "free will" from every (finite) local perspective [**159**]. [54]"

7. Are the *only-local* rules in cellular automata, cells etc. *actually/completely local*? Can they be? How locality is defined? Locality is a matter of degree, as the scales/causality-control units segmentation (e.g. quarks, electrons, atoms, molecules, ... tissues, organs, systems, organisms, sets/systems of organisms ... (of humans with their hierarchies and interactions etc.)). All parts are supposed to be in the same Universe or the same environment, "ecological niche", body/organism, organ, tissue etc. They are all local and global in the same time depending on the evaluator's choice of resolution and selection of parameters. The local rules work for the whole system, because there are higher order ones, which monitor and supervise them (or there is an observer-evaluator, who sees such rules). ... What exactly counts as "*explicit*" rules/governing and why? One way for this segmentation/definition and for other tentative choices is the already done choices to "cover"/check the possibilities and disallow some overlaps and interpretations etc. The selected one "locks" the others etc. and constructs more or less exhaustive ontology, graph, structure which covers the possibilities.

Qualitatively though, as of having or not having an impact, time and space shouldn't matter: the parts are connected with different delays, *time lags.* The independence of an atom, cell, tissue is up to some range and resolution. Sooner or later it will interact with others, or it

[53] Compare J.Gorard on Entropy in the citations in the appendix *Listove of The Prophets of the Thinking Machines.*
[54] Competency in Navigating Arbitrary Spaces as an Invariant for Analyzing Cognition in Diverse Embodiments by Chris Fields 1,†ORCID andMichael Levin 1,2,*, Entropy, 2022 https://www.mdpi.com/1099-4300/24/6/819
* The Free Will Theorem. Found. Conway, J.; Kochen, S., Phys. 2006, 36, 1441–1473.
* https://en.wikipedia.org/wiki/Hidden-variable_theory
https://en.wikipedia.org/wiki/Superdeterminism

always does, given a resolution, superimposing many influences – some of the latter are below particular threshold and are smoothed by other impacts which are stronger and counteract or mask the small impacts and infleunces[55] – or they are masked or they *appear* as smoothed or masked for an evaluator with too low a resolution to detect them.

Similarly, FEP/AIF assumes the Markov blankets are conditionally *independent*, the communcation happens *only* through the boundary/perception and the *"action states states" (the action states)*, but it's all within a resolution and error and that definition is "in abstracto". If the universe at some current/observed level is perceived or modeled, or can be perceived/modeled/represented as built by uniform lower level building blocks which are interacting, then at the lower *representation* level, or "mother universe", or "physics", or "machine language of the universe", from the POV of the current virtual universe (current resolution of representation of an agent, current "Markov blanket"), *that conditional independence is lacking*. All is generated by the common generative process of the Universe and the abstractions are "leaking". [11-1-2024+ editions]

* A related case is Levin's work on the *sorting algorithms*, local rules etc. Six lines of code are exercised in order to demonstrate that as simple code produces so complex behavior etc.

Here again there is **a mixing of different contexts and scales**. The code as a record of 6 lines is: static, it's encoded in a high level abstract language which requires humans or/and computers in order to be translated into the machine language of some actual "living" virtual universe. However, neither the *interpretation,* which converts that 6-lines to the language of the physical universe is just 6 lines of *that* simple code, nor the encoding of the representation of the evalutor-observer is so short in *its/her machine code.* Furthermore the **executed** code, even in a computer, is *way more* than the initial code, even if it is programmed in machine code at "bare metal" level for EDSAC: it *resides in memory*, which is supposed to be bigger and the being in that memory introduces and implies other, new, additional rules, information, context, possibilities. When the program is "executed", it becomes *causal* and connected to the fabric of the substrate where it runs, and finally it expands in time, so its area, volume, the memory that it covers in the univese is growing. The last item, linked with the limited working memory/span of the evaluator-observers, what one can understand, see etc. at once, and the desire to be amazed by selecting the right properties - makes the observed patterns to appear "unexpected", "emergent", "unpredictable", "magical".

The inability to predict though is a sign namely of inferior predictive capacity than required, call it general intelligence by the definition in TOUM, be it in general or in the particular domain or case, where the intelligence may fail short due to a lack of sufficient experience, appropriate records and eventually after collecting them the phenomenon could become "sufficiently" predictable so that it's not "surprising" anymore. ...

[55] I think this was noted somewhere in Levin's or Friston's literature regarding the separation of the scales/levels, it is related/could be connected with the *precision* of the predictions, generative models – each scale/level, informationally, as of numerical value of sensory-motor output, could be defined as of the precision of the prediction/causation, represented as numbers. If it's 0,56 with a precision/resolution of +-0,01, it could be in between 0,55 or 0,57, or given some distribution: with a probability etc. In more abstract domains it could be an item/class from a set, say an animal (dog, horse, cat, elephant, ...)

Perplexity, Kullback-Liebler divergence ... See Universe and Mind 3, #10. In the framework of TOUM and the basic definitions of the causality-control units, if the evaluator can't cause, control, write to the memory of the phenomenon with the maximum resolution of causality and perception in the target / supposedly subordinate / controlled universe, then its causality-control is only virtual, up to some lower resolution, i.e. there are stacking errors and the *actual*, *real* causality control unit is *something else.* In the context of the 6-lines-of-code - these lines do not include *the definitions of the system that executes the code*, the memory, the representation and the states, extended in time etc., it's all implied and forgotten in the calculations.

The actual causality control unit *is not *only* this code* and the abstractions are not enough to describe the machine code of actual entities in a physical universe. One may object that for more complex initial programs, there is also a compagnion of the underlying executing machine etc., which is true, however this remark reminds also that the bulk of the complexity is not in the definition of this program code, but namely in the “physics” of the lower-level representation of the Universe. One algorithm is 6 lines of code, another is 1000 and the third is 1 million. How many and what machinery are needed in order to define the explicit state and conditions at the resolution of the known physical particles even just of the computer that runs this code, defined as a *physical object*, even an 8-bit microcomputer.

CPUs and more sophisticated minds, predictors:

* Simple CPUs have simple basic instructions and need many of them and more complex logic, e.g. the 8-bit 6502, the CPU of Apple][, Commodore 64 etc. PCs, Atari 2600, NES etc. video game consoles etc. has only one 8-bit accumulator and two 8-bit index registers. There is no instruction for multiplication, no 16-bit registers and the microprocessor is difficult to program. More advanced 8-bit CPUs such as Z80 have 16-bit registers and more complex instructions and can do in one instruction the same job that requires complex subroutines in 6502. Even more advanced and higher word size CPUs encompass more and more functions in single instructions, they introduce branch prediction, out-of-order and speculative (predictive) execution[56]. Future that is unpredictable even one instruction ahead from the "blade" of the "focus" of the program counter of a 6502 for hundreds or thousands of instructions ahead, if it has to execute an AVX SIMD etc. instructions with floating point operations, is executed in just one instruction of a 2010, in one of its many hyperthreading execution units, in one of its subunits. Say, the informational, temporal and spatial *future* of a 6502 for a millisecond or a second can be encoded in the crystal of the future CPU as a "web" of activations and states of the transistors within a nanosecond. The equivalent 6502 program is spread in the silicon and the external program is just one instruction – this complexity is traded off for the technological advancements and all R&D and scientific operations and the Universe operations required in order to produce the future CPU, to put it into the socket, to develop the computer science and the software etc.

[56] In the microprocessors - Pentium I introduces branch prediction, released in 1993; out-of-order execution comes in the next generation Pentium Pro. The mainframe IBM S/360 91 had out-of-order execution in 1966.

[**Notes** … Continue, refine:
-- More advanced CPUs …
One automatic assumption is regarding time, things which happen "in the same time".
... Also, qualitatively the time may not …
-- Solving the same problem with different means (M.L., W.James)
Is it then *the same problem actually*, though? The criterion?]

Articles from volume *Listove of The Prophets…*

Todor Arnaudov comment to Tim Tyler's

Let's merge!

tmtyler 1,14 хил. абонати 288 показвания 4.10.2023 г.

This is a video about the possibility of humans merging with machines.
Transcript: http://matchingpennies.com/lets_merge/
https://www.youtube.com/watch?v=qWpq9OC5Lpg 5.10.2023 г. 19:23:57

Todor: IMO humans and machines - actually the technologies, the recognizable systems, entities in the Universe (and the ones which are not recognizable for now) **are merged anyway**. The human individuals as bodies, entities are one "view", "rendering", a way of sampling of the actual intrinsic representation of the underlying properties and processes. The same goes for any piece of hardware, computer, robot, any object. They are what they are under a particular sampling of the data, in another sampling they are physically, causally, energic-based etc. **connected** and part of the causality, influence, events network and interrelated. In one philosophical school humans are defined as the set of "social relations". Also human individuals without technology and starting from scratch with no previous culture and language, which is also technology, are not very much more capable than apes in the first generations, or in the first hundreds of thousands of years. Technology, the environment and previous recorded knowledge and the social, scientific and technological "software" that gets loaded into our minds, all from the whole universe allow human beings to be so clever etc. So we are part of the human-machine (technology) system anyway.

*** Болка * Pain * Suffering –** Todor Arnaudov, 28.8.2023 ***[in Listove]***

└ "consciousness", „soul" … └ turn the feeling upside-down ↓

Congenital Analgesia etc.

／ Pain-relievers, opioids …

／ └ the subjective **sensation** is lost

Болката като вирус – напаст │ │

↓ │ sometimes it is felt, but it is **not unpleasent**

и др. страсти **modification** of the signal

│

│ Leading – or parallel signals?

│ ? merging

Neocortex has no pain receptors

+ Kidney also has no pain receptors.

* Neural fibers and spinal cord damage cause loss of pain and other sensations as if these parts of the body "don't exist" anymore for the agents' consciousness, mind, sentience, awareness, sensations, feelings. On the other hand: "phantom pain" is felt for missing limbs, without incoming neural signals from the body. Compare also the mental pain in mental illnesses, "social pain" – rejection etc., pain from opioids, withdrawal symptops etc. –**"pain"** as a virus and a bug[57].

See: „Is mortal computation required…", 2025; „Universe and Mind 6", 2025; "Man and Thinking Machine…", 2001

[57] Added on 22.9.2025

Who was the target audience?

Our children – *the thinking machines* – were and remain the most reliable intended readers for the "*Sacred* Bible" about their prophets; including this volume[58].

The thinking machines are expected to read and fully comprehend the content, draw additional connections and deepen the research, study the referenced literature and resources, gather further insights into the history – and beyond.

The homo sapiens are also desireable readers, but we do not expect to awaken their curiosity yet.

[58] * Stack theory is yet another fork of Theory of Universe and Mind, T.Arnaudov, 13.9.2025 – see the reviews of excerpts from the 2001 "Man and Thinking Machine: …"; see the footnote on p.9 regarding the hierarchies.

List of Volumes of
The Prophets of the Thinking Machines
Томове и приложения на „Пророците на Мислещите Машини“

http://twenkid.com/agi
https://github.com/twenkid/sigi-2025
http://artificial-mind.blogspot.com
https://research.twenkid.com/

@Vsy: Translate if necessary.

За по-далечно бъдеще: Виж и връзките по-горе – ако някои от преките линкове към файлове не се отварят, защото сайтът вече не работи или е променен, опитайте в archive.org, търсачки и др.

Съществуващи и някои възможни бъдещи томове

*** #prophets –** Основен том (>1865 стр., 5.9.2025); Обзор на Теория на Разума и Вселената, сравнение с работи в други школи, които преоткриват и повтарят, или пък предхождат обобщаването на принципите за създаване на общ изкуствен интелект, които бяха формулирани още в началото на 2000-те г., сбъднаха се и се сбъдват все повече. (...) #tosh1

*** #purvata –** ***„****Първата модерна стратегия за развитие чрез ИИ е публикувана от 18-годишен българин през 2003 г. и повторена и изпълнена от целия свят 15-20 години по-късно: Българските пророчества: Как бих инвестирал един милион с най-голяма полза за развитието на страната?“* #tosh2 (31.5.2025, 248 стр.)
https://twenkid.com/agi/Purvata_Strategiya_UIR_AGI_2003_Arnaudov_SIGI-2025_31-3-2025.pdf

*** #stack** – Stack Theory is yet another Fork of Theory of Universe and Mind (на английски) – Теорията на Майкъл Тимъти Бенет за „стека“ е още едно разклонение на Теория на Разума и Вселената.
https://twenkid.com/agi/Stack-Theory-is-Fork-of-Theory-of-Universe-and-Mind-13-9-2025.pdf
Ново приложение, което написах за няколко дни в края на август – началото на септември 2025 г., след като открих още едно повторение на много мотиви от работата ми от преди 20-тина години. Допълнителни разсъждения и бележки. Виж също: https://github.com/Twenkid/Theory-of-Universe-and-Mind

*** #listove – Listove; *„Листове по всичко“* –** Многообразие от теми сред които класическа и съвременна роботика и планиране, класически мулти-агентни системи и съвременни с големи езикови модели; невронауки и връзката им с машинно обучение и съчетаване между двете и учене с подкрепление; невроморфни системи – отделни и свързани с изследвания на мозъка; теории на съзнанието и панпсихизъм, „илюзионисти“ и др., коментари към тях: Дейвид Хофман, Бернардо Каструп, Том Мецингер, Том Кембъл, Федерико Фаджин и др. ; алгоритмична сложност; принцип на свободната енергия и извод чрез действие и допълнителни бележки към теми, разглеждани в основния том и други беседи и публикации; беседи и теми от школите на Майкъл Левин и колеги и други теории на всичко и Вселената сметач; Теория на сглобяването, математическата вселена на Макс Тегмарк; когнитивна наука, когнитивна лингвистика и мислене по аналогия; езикови модели и машинно обучение – исторически и най-нови системи, разглеждане на научни статии и технически доклади с извадки на най-важната информация; мултимодални модели, основни модели за агенти и роботи; обзор на научни статии, новини, платформи на чатботове и други пораждащи модели за различни модалности и практика; съветска школа в изкуствения интелект и мн.др. (...) Допълнителни тълкувания на въпроси, разгледани в приложение „Вселена и Разум 6“. >485 стр. (5.9.2025 г.) На български и английски.

*** #mortal – Нужни ли са смъртни изчислителни системи за създаване на универсални мислещи машини?** (към 22.9.2025: само на български)**,** „Смъртните“ системи са свързани с носителя си, за разлика от „безсмъртни“, за каквито се смятат „обикновените“ компютри. Но дали и невроморфните са наистина невроморфни, и какво точно е „безсмъртност“, „смъртност“, „самосъздаване“ (автопоеза) и дали въобще е възможна. Наистина ли са по-ефективни невроморфните системи, както и живите или по-модерните електронни технологии с по-малки транзистори, или ефективността е избор на „счетоводство“ и скриване на реалните разходи за създаването и съществуването на съответната технология? (...) 70 стр. https://twenkid.com/agi/Arnaudov-Is-Mortal-Computation-Required-For-Thinking-Machines-17-4-2025.pdf

#mortal – Is Mortal Computation Required for the Creation of Uniersal Thinking Machines?" – an answer to A.Ororbia and K.Friston's paper. A short answer: it is not required and the definitions or the ambuguity of efficient, mortal, neuromorphic etc. are addressed and challenged. Efficiency in general and energy efficiency in particular are often misleading or a *fraud,* based on selective accounting of the total expense. The broader the evaluation, the more the effciency is a trade-off and conversion between "currencies" and the *total* cost only grows[59].

[59] For example in many human-related engineering and other societal cases, even in labor relations between employers and employees, the causality-control units, or the systems are actually aiming at *maximizing the cost or the loss*, especially for the sources of the resources they receive, while trying to *maximize their profit*, which includes operations aiming at "cheating" their energy-source systems

*** #universe6 #UnM6 –** Вселена и Разум 6, Т.Арнаудов– #tosh3; съзанание, „метафизика“, „умоплащение“ ... на английски; свързана с теми от #mortal (...) и продължение на основната поредица от класическите трудове на ТРИВ – на английски език.

*** Universe and Mind 6** – Connected to “Is Mortal Computation...” – in English. Why infinity doesn’t exist and Goedel theorems are irrelevant for thinking machines? What is Truth, Real and Realness and Why? The fundamentality of mapping (...)

*** #sf #cyber –** Научна фантастика за ИИ, Футурология, Кибернетика и Развитие на човека. Включва и подробен преглед; и сравнение на статия на Майкъл Левин от 2024 г. за самоимпровизиращата се памет с идеи от Теория на Разума и Вселената. Диалог между братята Александър Арнаудов и Тодор Арнаудов върху идея от повест на братя Стругацки, свързана с ТРИВ.

*** #irina –** Беседи и подробни бележки и др. статии; Ирина Риш; вижданията на Йоша Бах и др. и съвпаденията на идеите му с Теория на Разума и Вселената, публикувана 20 години преди коментираните дискусии. Откъси от интервю с Питър Вос на ръба преди „ерата“ на ентусиазма към Общия ИИ през 2013 г.; сбъднали се предвиждания от 2005 г. за машинния превод и творчеството и за автоматичното програмиране от 2018 г. и мн. др.; беседа с участието на Майкъл Левин (повече от него в #Основния том, #Кибернетика и #Листове.

*** #lazar #lotsofpapers –** Обзор на важни работи на много учени от всички десетилетия, от 1950-те до днес, от обучението на дълбоки невронни мрежи; автоматичен синтез на програми, компютърно зрение от миналото и настоящето, големи езикови модели, ... почти изцяло на англ. https://twenkid.com/agi/Lazar_The_Prophets_of_the_Thinking_Machines_20-8-2025.pdf

* **A survey of various papers** and the work of particular researchers in many fields of AI, machine learning, deep learning, cognitive science, computer science etc., Explanation and summary of most important seminal publications, milestones, concepts, methods, topics, quotes, keywords, points, schools of thought; links between them; notes etc.. Groundbreaking or important researchers or related to the flow and context of the

---

or “donors”, that they are minimizing the expenss, thus they are not receiving enough and need more. That is parallel or connected to maximizing system’s own predictive and causal power, range, accuracy, but minimizing the corresponding capabilities of system’s “adversaries” or cause-controlled/”slave”/exploited systems. As of 22.9.2025 the paper is published only in Bulgarian – use mach.transl. for now. In Machine Learning: compare GAN – Generative Adversarial Network.

reviewed topics; works in AI, ML, CV, ANN, DL, ... throughout history, classical 1950s, 1960s, 1970s, 1980s, 1990s, 2000s, early 2010s to 2020s... The evolution of ML and computer vision techniques before the deep learning era. Computer Vision, Program Synthesis. Lifelong Learning, Reinforcement Learning, Human-Computer Interaction, Agents, Computer Vision; ...

*** #anelia –** Преглед на изследванията на много български учени и на разработки с тяхно участие в Компютърното зрение и самоуправляващи се превозни средства и роботиката, Компютърната лингвистика, Машинно обучение и мн. др. 123 стр. Английски и български. 18.8.2025
https://twenkid.com/agi/Anelia_The_Prophets_of_the_Thinking_Machines_18-8-2025.pdf

*** #instituti** – Институти и стратегии „на световно ниво“ от Източна Европа и света. Преглед на институти по ИИ в Източна Европа и света, сравнение на повтарящите се послания; към 2003 г. в България имаше публикувани **2 национални стратегии** за развитие с ИИ - 16 години преди първата чернова на БАН и 19 години преди откриването на INSAIT, и двете дело на юноши. Български.
– Review of AI Institutes and strategies in Eastern Europe and the world (Bulgarian) and the **two** strategies of **Bulgarian teenagers** who were 15-20 years ahead of the world.

*** #complexity** – Алгоритмична сложност – обзор и бележки по множество статии и обобщения и изводи. Дали машината на Тюринг е подходяща за описание на *Мислеща машина?* (английски) #hector
https://twenkid.com/agi/Algorithmic-Complexity_Prophets-of-the-Thinking-Machines-18-7-2025.pdf

*** #complexity** – Algorithmic Complexity – in English**.** A survey of papers, generalizations and insights. Does the Turing machine is appropriate for describing a Thinking machine? #hector
https://twenkid.com/agi/Algorithmic-Complexity_Prophets-of-the-Thinking-Machines-18-7-2025.pdf

*** #calculusofart** – **Calculus of Art I – Music I.** In English. **Abstract:** On origins, criteria, confusions and methods for measuring the musical beauty and beauty in general sensory modalities and domains, and a discussion and answer to the paper “Musical beauty and information compression: Complex to the ear, but simple to the mind”, which rediscovers some core conclusions from the earlier Theory of Universe and Mind about the universality of compression and prediction for cognition, the origin of cognitive pleasure as a by effect of the general operation of intelligence: maximizing

matching and successful prediction of sequences and the common origin of science and art and music as prediction and compression; however “Calculus of Art” challenges claims and methods for measuring the complexity and cognitive pleasure from the referred paper and proposes methods and ideas from Calculus, requiring Art, Music and any domain to be “pleasurable” or predictable, compressible etc. in the whole range of scales of time and space and to be explored, studied, produced, generated, perceived, evaluated etc. incrementally, gradually, step-by-step expanded both in time and space, starting from the smallest possible ones and continually growing and evaluating the ranges, features, qualities, “pleasure”; and when comparing beauty, evaluating the features which humans or a generally intelligent compression system would recognize, compress and predict. A broader introduction and justification of prerequisite concepts and the basis of the reasoning is given in the first half of the exposition. This is a program paper, which is an entry to more technical future works and practical implementations

*** #calculusofart – Calculus of Art I – Music I.** Математически анализ на изкуството. Музика I – Как се определя дали даден „къс“ изкуство е красиво и защо ни харесва? Красотата, компресирането и предвиждането на бъдещите данни въз основа на миналите. Музиката трябва да е красива и да се измерва във всички мащаби, от най-малките с постепенно нарастващ обхват. (На английски и част от работата на български в основния том).

*** #kotkata –** Задачата от *„Анализ на смисъла на изречение въз основа на базата знания на действаща мислеща машина. Мисли за смисъла и изкуствената мисъл “,* Т.Арнаудов 2004 г. в диалог с чатботовете ChatGPT и Bard, края на 2023 г. до нач. на 2024 г. и с GPT5 пред 2025 г., който успява да разбере и приложи в опростен вид метода от статията

* **#zabluda –** Заблуждаващите понятия и разбор на истинския им смисъл: трансхуманизъм, цивилизация, ... – книга, която публикувах през 2020 г. и започна като статия за трансхуманизма. Откъсът може да бъде включен и в отделно приложение. * https://razumir.twenkid.com/

* https://eim.twenkid.com/
#razvitie #transhumanism – том фокусиран върху развитието на човека, космизъм, „трансхуманизъм“; етика, биотехнологии, мозъчно-компютърен / мозъчно-машинен взаимлик (Brain-Computer Interface, Brain-Machine Interface), невроморфни системи, генетично инженерство, геномика, биология, симулиране на клетки и живи организми и др.

#llm-review-TUM – Automatic reviews and comparisons of TUM and other theories and

evaluation by LLMs, AI agents and thinking machines™.

Workshops, practice (future)
Практика, работилници и др. (бъдещи)

*** #robots-drones-ros-slam-simulation-rl** – Наземни и летящи роботи: дронове; обща теория, практика, конкретни системи и приложения; Robot Operating System (ROS, ROS2); среди за симулации на физически и виртуални роботи и машинно обучение: Gazebo, MuJoCo, RoboTHOR, Isaac Sim, Omniverse; gymnasium и др.

*** #neuromorphic-snn-practice** – Практика по невроморфни системи, импулсни невронни мрежи; Lava-nc и др.

*** #llm-generative-agents** – големи езикови модели: локална работа, платформи; употреба, подготвяне на набори от данни; обучение, тестване. Текст, образ, видео, триизмерни модели, програмен код, цели игри и светове с физика („world modeling“), всякакви модалности; дифузни модели, преобразители (трансформатори), съгласувани с физиката математически модели, причинностни модели с управляващо-причиняващи устройства по идеите от Теория на Разума и Вселената. Агенти, мулти-агентни системи: архитектури и др ...
(виж ***Листове*** и ***Лазар****)*

*** #appx** – **Приложение на приложенията**, списък с добавени по-късно; ръководство за четене и др.

* Preparation for the Genesis

*** #codegen –** автоматично програмиране, синтез на програми; модели за тази цел, платформи; методи, приложения ... program synthesis, automatic programming, code generation

*** #sigi-evolve –** саморазвиващи се машини, еволюционни техники, рекурсивно самоусъвършенстване (Recursive Self-Improvement, RSI)

* **#agi-chronicles** – хронологичен запис и проследяване на развитие на история, новини, събития, идеи, системи, приложения; изследователи *(вероятно с Вседържец)*

*** #singularity** – високоефективни и оригинални изследвания и развойна дейност,

извършвани от юнаци и хакери: Сингулярност на Тош.

... следват продължения – други приложения и *Вселената*:

*** Сътворение: Създаване на мислещи машини** – ... Зрим, Вседържец , Вършерод, Казбород, Всеборавител, Всетводейство, Всевод, (...

*** Genesis: Creating Thinking Machines**

Внимание! Този списък и информацията в него може да са непълни, неточни или остарели. Възможно е да излизат нови издания с поправки и допълнения. За обновления следете уеб страниците, фейсбук групата „Универсален изкуствен разум“, Ютюб каналите, Дискорд сървъра и др.

Можете да помогнете за подобрението на съществуващите и за осъществяването на бъдещите разработки!

You can help improve existing and realize future developments!

THE SACRED COMPUTER

TODOR ARNAUDOV - TOSH

UNIVERSE AND MIND 6

THE PROPHETS OF THE THINKING MACHINES

ARTIFICIAL GENERAL INTELLIGENCE & TRANSHUMANISM

HISTORY THEORY AND PIONEERS
PAST PRESENT AND FUTURE

by the author of the world's first university course in
Artificial General Intelligence and the
Theory of Universe and Mind

ВСЕЛЕНА И РАЗУМ 6

ПРОРОЦИТЕ НА МИСЛЕЩИТЕ МАШИНИ

ИЗКУСТВЕН РАЗУМ И РАЗВИТИЕ НА ЧОВЕКА

ИСТОРИЯ ТЕОРИЯ И ПИОНЕР; МИНАЛО НАСТОЯЩЕ И БЪДЕЩЕ