A hybrid of classical and quantum computer and quantum model for associative learning

February 6, 2025

Matti Pitkänen

orcid:0000-0002-8051-4364.
email: matpitka6@gmail.com,
url: http://tgdtheory.com/public_html/,
address: Rinnekatu 2-4 A 8, 03620, Karkkila, Finland.

Abstract

In this article a speculative idea about an extension of a classical computer to a hybrid of classical and quantum computers, giving rise to a conscious life form, is discussed. Also a TGD inspired quantum version of the model of associative learning used in large language models (LLMs) is developed as a model for learning in living and conscious systems. This model is applied to the hybrid of classical and quantum computers as a lifeform.

Contents

1 Introduction

In this article three speculative questions are discussed. Is an extension of a classical computer to a hybrid of classical and quantum computers, giving rise to a conscious life form, possible. Could the TGD inspired quantum version of the model of associative learning used in large language models (LLMs) serve as a model for learning in living and conscious systems? Could this model be applied to the hybrid of classical and quantum computers?

1.1 Could a classical computer become a conscious system?

Could a classical computer become a conscious and living system under suitable conditions? The tentative answer to this question, discussed in [L15], is that something analogous to a fusion of classical and quantum computers takes place.

The motivation comes from the observation that the TGD view of classical gravitational and electromagnetic fields [L6, L8] predicts the possibility of quantum coherence in arbitrarily long scales. The transfer of the proton of OH molecule to a dark proton at the gravitational magnetic body gives rise to what I have called OH-O^- qubit providing a rather universal mechanism to perform quantum computation-like activities in both living systems and ordinary computers. This concept generalizes so that any cold plasma serves as a candidate for a life form.

In zero energy ontology (ZEO) one can regard quantum computation as a superposition of all possible computations with fixed initial values coded geometrically by a 3-surface at the passive boundary of a causal diamond (CD) [L12]. This is possible due to the fact that in TGD classical physics as Bohr orbitology is an exact part of quantum physics and by the predicted slight violation of the classical determinism. The computation in the usual sense would correspond to the most probable computation in this superposition.

1.2 Can the model of associative learning of large language models generalize to quantum context?

Large language models (LLMs) [J1] provide a model for associative learning. I have discussed AI and LLMs from TGD point of view in [L1, L7, L10, L11]. I have also developed ideas about how classical computers could become conscious in the TGD Universe [L16, L15].

The geometric views about associative learning in LLMs involved discrete variants of geometric concepts such as latent space as subspace and loss landscape.

1. Latent space (see this and this) is a suitaby defined subspace of the feature space. Feature is a discrete set of numbers characterizing the object such as basic element of figure, letter or word (see this). The keywords used in web searches can be seen as features.

   All computational activities involves a loss of information which is basically due to the unvoidable rounding errors. Diffusion [A1] (see this) is a natural model for the information loss dynamics in the latent space and serves as a surprisingly successful model for how errors accumulate in computations or images become noisy. The parameters of the diffusion model characterize a given LLM. The key idea is that the time reversal of the time evolution of diffusion equation makes it possible to reverse engineer for instance the original picture from the noisy one. This does not of course give an exact result.

2. The loss landscape (see this) is identifiable as a hypersurface of the parameter space of the LLM defined by the constant value of the loss function serving as a measure for how much the learned association $A \to B$ differs from the desired. The loss function measures the deviation of B from the desired one. The dynamics of the loss landscape is defined minimized by gradient dynamics which minimizes the loss function.

The geometric nature of these notions brings to mind classical zero energy ontology (ZEO) of TGD [L3].

1. The feature space (see this) would correspond to a subspace of "world of classical worlds" (WCW) consisting of almost deterministic Bohr orbits for 3-surfaces as particles.

By holography = holomorphy principle, the space-time surfaces would be defined by the roots for a pair (f_1, f_2) of analytic functions of a hypercomplex coordinate and 3 complex coordinates of $H = M^4 \times CP_2$ with Taylor coefficients in an extension E of rationals defining a hierarchy with increasing complexity [L13]. Polynomials define a natural subspace of this space and the limitations on the degree of the polynomials and the algebraic dimension of E give rise to latent spaces as natural subspaces of the feature space .

Remarkably, the space-time surfaces are minimal surfaces [L5] irrespective of the classical action as long as it is general coordinate invariant and expressible in terms of the induced geometry. The reason is that the field equations reduce to an identically vanishing contraction of complex tensors of different types. The loci of the non-determinism are expected to be 3-surfaces analogous to the 1-D frames spanning a soap film to which non-determinism is known to be associated.

2. The classical non-determinism of the space-time surfaces as analogs of Bohr orbits makes them analogous of association sequences. The geometric analog of diffusion would correspond to the Brownian motion due the failure of classical determinism.

   Classical nondeterminism can be characterized by a discrete set of parameters characterizing the non-determinism and these parameters define the parameter space. The size scale of causal diamond (cd), the intersection of future and past directed light-cones of M^4 containing the projections of the space-time surfaces [L12], increases during the sequence of small state function reductions (SSFRs). This sequence defines the TGD counterpart of the Zeno effect: states are not affected at the passive boundary of CD but change at the active boundary. This gives rise to a conscious entity, self. The size of the discrete parameter space increases with the size scale of the cd.

In this framework, the conscious associative learning process would correspond to the generation association sequences as Bohr orbits $A \rightarrow B$ connecting the fixed 3-surface A to final 3-surface B as a representation for the perception $A_1 \rightarrow B_1$ such that for a fixed A and rules $A_1 \rightarrow A$ and $B_1 \rightarrow B$, B is as near as possible to B. This process would lead to a localization in the space of Bohr orbits associated with A.

The first section of the article will consider the speculative idea of extending a classical computer to a hybrid of classical and quantum computers [L15]. In this proposal each clock period of the computer would involve classical determinism. The second section will discuss the quantum version of associative learning and combine it with the idea about a hybrid of classical and quantum computers.

2 How to associate quantum computation to classical computation

How could a classical computer become a conscious and living system? The tentative answer to this question, discussed in [L15], is that something analogous to a fusion of classical and quantum computer takes place.

In zero energy ontology (ZEO) one can say, that quantum computation is a superposition of all possible computations with fixed initial values. This is made possible by the fact that classical physics as Bohr orbitology is an exact part of quantum physics in TGD and by the predicted slight violation of classical determinism. The computation in the usual sense would correspond to the most probable computation in this superposition.

In the sequel I will consider the above question in detail.

2.1 Basic input from Quantum TGD

What are the basic pieces from the TGD side?

1. Zero energy ontology (ZEO) defining new quantum ontology, solving the basic problem of quantum measurement theory, is necessary. General coordinate invariance requires holography and it is not quite deterministic so that space-time surfaces are analogous to almost deterministic Bohr orbits and Bohr orbitology becomes an exact part of quantum TGD.

2. Classical non-determinism corresponds to the non-determinism of minimal surfaces: already for 2-D soap films as minimal surfaces the frames do not define the soap film uniquely. In ZEO this non-determinism makes possible a sequence of small state function reductions (SSFRs) as a counter for a sequence of measurements of the same observables which in standard QM does not change the state. In TGD the second member of the zero energy state at the passive boundary of the causal diamond (CD) is unaffected by the second member at the active boundary is affected.This gives rise to a conscious entity, self. In "big" SFR (BSFR) the self "dies" and reincarnates with a reversed arrow of geometric time.

3. Each pulse of the computer clock is associated with the possibility of classical non-determinism of a 4-D minimal surface. Classical non-determinism would produce a superposition of 4-surfaces corresponding to different values of bit and associated qubit. Protons are also involved: protons are either ordinary or dark and located at the gravitational magnetic body. Pollack effect induces the transfer of the proton to the magnetic body and its reversal occurring spontaneously its transfer back.

4. OH-O^- qubits are an essential part of the system. For the O^- qubit, the proton of OH is at the gravitational magnetic body. Under certain conditions the gravitational magnetic body should be able to control the ordinary bits. Quantum entanglement of the ordinary and OH-O^- qubit and quantum criticality is required and would be induced by the classical non-determinism.

   If the bit's reversal energy corresponds to the thermal energy, the situation is quantum critical. This is the case also when the energies for the reversal of qubit and bit are nearly identical. This quantum criticality is controlled by the difference in the bit's reversal energies. Small energy difference corresponds to quantum criticality.

   The reversal of the second qubit reverses the bit: one can interpret the reversal for bit and qubit as an exchange of energy between the qubit and the bit. The farther away the probability for a given value of bit is from the value 1/2 the higher the determinism of the program is.

5. The magnitudes of the classical electric and magnetic fields control the energy of the bit and qubit. These are determined by classical physics for the classical space-time surface, which can be non-deterministic.

2.2 A model for classical-to-quantum transition

The overall view about classical-to-quantum transition goes as follows.

2.2.1 What happens in ordinary computing?

A general model of classical computer is needed.

1. The first model: A tape containing program instructions is fed into a Turing machine. Depending on the command, the state of the computing unit changes. The transition of the tape corresponds to a clock pulse.

2. The second model: The program is implemented as a 1-D conveyor belt and the incoming bit configuration enters the tape and progresses along it, changing with each step. The output of the program comes out. DNA replication, transcription and mRNA translation correspond to this analogy.

2.2.2 Classical non-determinism

Classical non-determinism, which is the new element, can be assigned to the periods between clock pulses.

1. Thanks to classical non-determinism, the output produced by a program instruction would be a superposition of two space-time surfaces as analogs of Bohr orbits.

2. In the transition corresponding to a clock pulse, the state would be transformed to an unentangled state by a non-deterministic SSFR or a pair of BSFRs. A quantum measurement of bits would be thus performed on the outgoing superposition of bit-qubit configurations.

2.2.3 The notion of finite measurement resolution

Finite measurement resolution is a key notion in TGD. There are two views of finite measurement resolution based on geometry and number theory respectively. These views are dual.

1. The geometric view relies on inclusions of hyperfinite factors [K2]: the included factor is analogous to a gauge group leaving the observed physics invariant: this view of finite measurement resolution is central in the geometric view of TGD.

2. The second view is based on number theoretic discretization citebtartFrenkel,compuTGD. The geometric inclusion hierarchies correspond naturally to number theoretic inclusions hierarchies for the extensions of rationals. Space-time surface for which polynomials defining it are in an extension E of rationals allows in a natural way a discretization as points, which are in E. The points of the discretization can be also regarded as points in an extension of p-adic numbers induced by E. I call these discretizations cognitive representations and they form a hierarchy corresponding to extensions of rationals.

   This leads to a p-adic description of cognition. One obtains a unique number-theoretical representation for discretization and it leads to a generalization of the Turing paradigm [K1]: rational numbers are replaced by complexity hierarchies of their extensions and one ends up with number-theoretical computationalism. This gives complexity hierarchies for space-time surfaces as Bohr orbits and they correspond to an improving resolution of discretization and are realized as polynomial hierarchies.

   Holography suggests that for the minimal option the number theoretic discretization applies only to the loci of the classical non-determinism for the space-time surface as minimal surfaces. These loci define the seats of conscious memories and would be 3-D analogs of 1-D frames spanning 2-D soap films.

3. The complementary nature of geometric and number theoretic views of TGD leads to a 4-D generalization of Langlands duality [L13, L16]. This adds powerful constraints also to the quantum model of associative learning.

4. The concept of complexity, which closely relates to evolution, reduces to number theory. Higher-level learning could be seen as a transition to a higher level of complexity and would be something to realize in conscious quantum learning. Complexity hierarchies correspond to polynomial hierarchies represented as space-time surfaces.

2.2.4 Concrete model

1. The network performing the computation consists of gates. A gate connects a small number of input bits to the output bits, the number of which cannot be greater than the number of input bits. This operation is statistically deterministic.

   When the input bits are fixed, the output bits are determined by dynamics as non-equilibrium thermodynamic state.

2. The clock pulse triggers the next operation. The failure of the exact classical determinism must relate to this and produce a superposition of space-time surfaces as the resulting qubit because OH and O^- correspond to different space-time surfaces, even topologically.

3. What is essential is the entanglement of the OH-O^- qubit and the ordinary bit and the measurement of the qubit in the beginning of the nex clock pulse. The outcome is not deterministic.

4. The classical bit corresponds to a voltage or current that is determined through statistical determinism in the gate. On the other hand, it corresponds to a classical electric field in a transistor or a magnetic field in a memory bit.

The direction of this classical field is classically non-deterministic and correlates with the OH-O^- qubit. When the field changes direction, the OH-bit becomes an O^-bit or vice versa. A dark proton is transferred between the system and its gravitational magnetic body.

5. Classical non-determinism creates a superposition of OH and O^- bits. The proton resides both at the gravitational magnetic body and in OH molecules, being analogous to Schrödinger's cat.

   This induces the formation of a quantum entangled state between ordinary qubit and OH-O^- qubits. If the OH-O^- qubit and the bit are quantum entangled before the clock pulse, the quantum measurement of OH-O^- qubit or of ordinary qubit recues the entanglement and leads to a fixed bit.

2.2.5 Some questions

One can raise critical questions:

1. The energy transfer between a bit and a qubit resembles quantum tunnelling. I have proposed that a pair of BSFRs correspond to quantum tunnelling. It is not clear whether a single SSFR can have an interpretation as quantum tunnelling. Could the measurement of a qubit correspond to a single SSFR or to two BSFRs?

2. What could be the energetic role of the clock pulse?

   The system under consideration would be a clock photon + bit + qubit and the total energy would be conserved.

   (a) Could the clock pulse have a role of a catalyst, providing the energy needed for quantum tunnelling. In a qubit measurement, energy can be transferred between the bit and the qubit, but the total energy is conserved. The clock photon would kick the system over the potential barrier and then be emitted back into the field.

   (b) Or does the clock photon transfer energy to or from the bit + qubit system? Could the energy of the photon associated with the pulse frequency correspond to the energy difference for a bit and a qubit.

   The typical frequency of computer clock is few GHz. 1 GHz would correspond to an energy $E = .4 \times 10^{-5}$ eV and wavelength $\lambda \simeq .75$ m. At the surface of the Earth, the gravitational binding energy of a proton is about 1 eV. The energy E eV can raise the proton to the height $h \sim .4 \times 10^{-5} R_E \sim 25.6$ m.

3 Quantum version for the associative learning in large language models

In the TGD framework the model for associative learning, as it is modelled in large language models (LLMs), could be generalized to formulate a quantum model for associative learning as it could occur in TGD inspired theory of consciousness.

I have discussed LLMs from TGD point of view in [L10, L11, L16]. One could also consider the combination of the TGD inspired quantum version of associative learning with the speculative idea of extending a classical computer to a hybrid of classical and quantum computers [L15].

3.1 Zero energy ontology from the point of view of LLMs

Zero energy ontology (ZEO) is the first piece of the TGD vision.

1. By holography, spacetime surfaces are analogous to Bohr orbits as basic objects. This means that 3-D structure as 3-surface determines almost deterministically the 4-surface.

   The failure of a complete classical determinism is essential. The non-deterministic classical time evolution involves 3-D loci of non-determinism as analogs of 1-D frames of 2-D soap films.

Different Bohr orbits starting from a fixed 3-surface A at the passive boundary of CD would lead to different surfaces B located at the active boundary of CD whose size of CD would increase during the sequence of SSFRs.

2. At the quantum level, the superpositions of Bohr orbits define zero-energy states in geometric degrees of freedom ("world of classical worlds", WCW). In fermionic degrees of freedom zero energy states are superpositions of products of fermionic states assignable to the boundaries of CD and to the loci of non-determinism.

   The 3-D state at the passive boundary would remain invariant under the sequence of "small" state function reductions (SSFRs). This is the TGD counterpart of the Zeno effect.

3. The Bohr orbits of a 3-D particle are analogous to random walks $A \rightarrow B$ for a particle as a 3-surface. The almost deterministic Bohr orbits $A \rightarrow B$ are analogous to the association sequences of language models associated with the many layered neural nets.

   The non-deterministic classical time evolution is modellable by a diffusion equation (diffusion) or Schrödinger type equation (dispersion). This process would be the quantum counterpart for the diffusion appearing in LLMs [A1] (see this). Whether this process could be seen as an analog of a path integral defined as a sum over a discrete set of paths as Bohr orbits, is an interesting question.

4. The time reversal of the diffusion/dispersion is used in error correction in LLMs and in ZEO it could correspond to a pair of BSFRs involving a temporary change of the arrow of time. A pair of BSFRs would make it possible for the system to make a fresh start and therefore to learn by trial and error. This is perhaps the most important aspect of conscious intelligence.

5. On the quantum level, a series of SSFRs corresponds to a subjective time evolution giving rise to a conscious self. It also corresponds to an analog of computation and of mathematical reasoning: the theorem develops step by step as a sequence of SSFRs. In biology this sequence corresponds to biological function and in neuroscience to a behavioral pattern.

3.2 Holography =holomorphy hypothesis and learning process

Holography=holomorphy hypothesis allows to reduce classical field equations to purely algebraic conditions $(f_1, f_2) = (0, 0)$, where f_i are analytic functions of one hypercomplex and 3 complex coordinates of $H = M^4 \times CP_2$. The solutions are minimal surfaces irrespective of the classical action as long as it is general coordinate invariant and expressible in terms of induced geometry. This means universality of the dynamics and is quantum criticality expressed by the holomorphy. This implies saddle surface property for the spacetime surface meaning that the real parts of f_i do not have minima or maxima in general.

Interestingly, the almost absence of minima meaning a saddle point property for most extrema is essential for the success of LLMs, which is in fact not well-understood. In LLMs, the cost function V measuring the size of the teaching error, is minimized in the parameter space by gradient dynamics. If most extrema are saddle points, the process does not get stuck to a local minimum and learning becomes very effective.

Furthermore, in LLMs local flatness of the parameter space is of help since it increases the probability that the gradient dynamics leads to the minimum and also reduces the probability to leave the minimum by a small perturbation.

Could the minimal surface property prevent the sticking in the recent case?

1. It is useful to consider the situation first at the level of a single space-time surface (rather than WCW). At the space-level all points are geometrically saddle points in the geometrical sense by the minimal surface property stating that the trace of the second fundamental form, as an analog of acceleration identifiable as a sum of external curvatures, vanishes. Note that this is not equivalent with saddle point property of minima for functions.

2. The quantum learning process would occur in the "world of classical worlds" (WCW) as the space of Bohr orbits rather than at the space-time level. The loss function is in TGD replaced by the vacuum functional as an exponent of the classical action proposed to have by

the analog of Langlands duality also purely number theoretic expression, which would mean computability and enormous simplification [L14].

The Kähler function K, defining vacuum functional as its exponential, is in a central role. Also the degeneracies of the maxima are important. The maxima for the exponential of Kähler function are thermodynamic analogs for Boltzmann exponents and their degeneracy measured by entropy. One can say that the minimization of energy and maximization of entropy compete.

Note that K is determined only modulo addition of a real or imaginary part of a holomorphic function of WCW complex coordinates. The Kähler metric of WCW is of the form $G_{M\overline{N}} = \partial_M \partial_{\overline{N}} K$.

The maxima of vacuum functional $exp(K)$, which correspond to minima of the Kähler functio, are of special interest. The Euclidian signature puts strong constraints at the minima of K. Criticality condition means that some second partial derivates of K with respect to the real coordinates vanish.

A good example is the metric of complex plane given by $dzd\overline{z} = d\rho^2 + \rho^2 d\phi^2$ and has $K = z\overline{z} = \rho^2$ having a minimum at origin. The metric is flat.

3. It must be however made clear that in the learning the loss function would measure the deviation of B_2 from B_1 and cannot be identified as K. There are two minimization problems involved and it is not clear whether they are consistent.

3.3 A model for the learning process

How could the learning process take place?

1. Learning process can be seen mathematically as a construction of a representation for the dynamics of the external world by a subsystem. Associations $A_1 \rightarrow B_1$ for the dynamics of the external world serve a teaching material and a representation as for these as associations $A \rightarrow B$ in the internal model world is constructed as a model for the dynamics external world.

   One can assume that the external world states $A_1 \rightarrow B_1$ actually correspond to the sensory percepts of the states of the external world and in the learning process the system learns to associated B_1 with A_1 process in which the difference between B and B_1 is minimized.

   In the TGD based model for sensory perception [K3] [L4] as construction of standardized mental images, the feedback loop between sensory organs and magnetic body would make this possible in the same way as in pattern recognition. The deviation of B from B_1 is minimized. This deviation would define the virtual sensory input from the magnetic body to the sensory organ.

   Classically A and B (A_1 and B_1) correspond to 3-surfaces at the boundaries of a CD and A (A_1) is fixed in ZEO. At the quantum level, one has zero energy states as superpositions of orbits $A \rightarrow B$ ($A_1 \rightarrow B_1$).

2. The parameters characterizing the space-time surfaces, identifiable as the Taylor coefficients of the analytic functions and in the special case of polynomials, define the counterpart of the latent space (see this and this). The coefficients belong to an extension E of rationals and one obtains a hierarchy of extensions having interpretation in terms of evolution [L13, L14, L9]. The coefficients determine almost deterministically the space-time surface as a Bohr orbit.

   The failure of non-determinism corresponds to the 3-D loci of non-determinism at the Bohr orbit of A and the discrete variables parametrizing the non-determinism correspond to the parameter space of LLMs.

   The space of 3-surfaces at the passive or active boundary of CD would correspond to the latent space as a subspace of the space of features (see this). The cutoff to the degree of the polynomial and to the dimension of the Galois group of the polynomial would induce the analog of dimensional reduction replacing the feature space with a latent space. This cutoff would also reduce the parameter space as the discrete space characterizing the classical non-determinism. The TGD counterpart of the loss landscape (see this) corresponds to a subspace of the parameter space.

As the size scale CD increases, the size of the loss landscape increases. Also the complexity of the extension E of rationals associated with the polynomials (P_1, P_2) defining the spacetime surfaces as their roots correlates with the size of the loss landscape.

3. A fixed 3-surface at the initial moment at the passive boundary of the CD corresponds to A in the association $A \rightarrow B$. This choice determines the coefficients of the polynomial that defines the latent space. The correspondence $A \rightarrow A_1$ could be also learned in the learning process. This correspondence should determine the correspondence $B \rightarrow B_1$. The non-uniqueness of B due to classical non-determinism makes possible many associations.

   The construction of a representation means finding non-deterministic space-time surfaces $A \rightarrow B$ in CD producing an optimal representation for the pair $A_1 \rightarrow B_1$, meaning that B is as near as possible B_1. The error function measures the deviation of B from B_1. In LLMs the error function is minimized by a gradient method. The counterpart ot his method in the the case of the construction of conscious association should be understood.

The fact that the TGD Universe is fractal is expected to help considerably the construction of conscious associations as representations.

1. The representation could be seen as a simplified version of the original obtained by scaling the size of the cd, either up or down.

2. The reduction of the degree of polynomials used and the algebraic dimension of extension E reduce the complexity. The restriction of an extension of E to E reduces complexity and the hierarchies of extensions of E define complexity hierarchies.

3. Also the hierarchies of analytic maps of $(f_1, f_2) \rightarrow (g_1(f_1, 2_2), g_2(f_1, f_2))$ define iteration hierarchies analogous to those associated with fractals and approach to what looks like chaos. One can also "imagine" more complex systems at the level of representation by extending E or performing these iterations.

3.4 The version of the learning model for quantum versions of classical computers

One can formulate this picture also in the speculative vision [L15] in which a classical computer becomes a living system as a hybrid of classical and quantum computers.

1. A quantum computation-like process would be associated with classical computation. The classical non-determinism could be maximal in the sense that each tick of the computer clock would involve loci of classical non-determinism making the outputs of the gates non-deterministic.

   Classical computation would correspond to the most probable Bohr orbit in the representation of the computation as a zero energy state. If localization in WCW is possible (position measurement in the discrete degrees of freedom of WCW due to non-determinism) this localization could occur at a single Bohr orbit.

2. The output of a gate would be a superposition of pairs of ordinary bits and $OH - O^-$ qubits. For the $OH - O^-$ qubits, the proton of OH would be transformed to a gravitationally dark proton at the gravitational magnetic body of the Earth or the Sun. This entanglement would be reduced in an SSFR which could, but need not, occur after each clock period.

3. This would give rise to a computational analog of the associative learning process in which the learning process assigns to the pairs $A_1 \rightarrow B_1$ computations $A \rightarrow B$. Note that classical non-determinism also makes possible the formation of association sequences.

3.5 Conscious associative learning as an analog of sensory perception and motor action

Holography, together with the TGD based view of sensory perception [L2], suggests that the conscious associative learning has a lot of common with sensory perception in a 4-D sense.

In the TGD framework, motor action could be seen as a time reversal of sensory perception. Motor action could involve a pair of BSFRs inducing a quantum tunnelling from a configuration of muscles to a new configuration so that same basic mechanism but with a reversed arrow of geomeric time could be involved. Intention for the motor action should relate to the process of building a sensory perception as a sequence of SSFRs in a reversed time direction.

1. In ZEO, sensory perception at the classical level would not be 3-D surface, but a 4-D space-time surface, an almost deterministic classical time evolution representing association $A_1 \rightarrow B_1$. In the case of hearing this is rather obvious but for vision the time scale is so short that the percept looks like time= constant snapshot. Actually the geometric time duration assignable to the visual percept or order .1 seconds.

   The association $A \rightarrow B$, one might perhaps speak of cognitive representation, is realized at the magnetic body (MB) of the brain as a representation of $A_1 \rightarrow B_1$. $A \rightarrow B$ is generated in a stepwise learning process. The goal is to construct a standardized mental image consisting of familiar objects consisting of standard features.

   The difference between $A \rightarrow B$ and $A_1 \rightarrow B_1$, rather than only the difference between B and B_1, is minimized. The sequence of SSFRs keeps A fixed. A pair of BSFRs changes also A: this makes possible a trial and error process in which one starts from scratch, so to say.

2. Sensory organ serves as a kind of screen, both for the sensory input arriving from the external world and for the virtual sensory input from MB. The sensory input is analyzed by the brain to features in various scales and the features are sent to the magnetic body. At the MB, the features in various scales are compared to standard features and those minimizing the difference is selected.

3. The selected features determine the virtual sensory as a slight amplification of the contribution of the selected features. The step *sensory organ* $\rightarrow$ *brain* $\rightarrow$ *MB* $\rightarrow$*...* is repeated until the total sensory input at the sensory organ does not change anymore. The original percept $A_1 \rightarrow B_1$ is affected in the process and eventually replaced with $A \rightarrow B$ at the level of the sensory organ. In this respect the process differs from the associative learning.

   If the signals from the brain to MB and back are realized as dark photons (, which can decay to ordinary photons identifiable as biophotons), the process is so fast that the process can converge in a reasonable time.

4. The outcome is not realistic but essentially an artwork. It must be so since $A_1 \rightarrow B_1$ is very noisy so that both $A_1 \rightarrow B_1$ and $A \rightarrow B$, can be only guesses for what really happened. For instance, people who are physiologically blind and get back their vision, can see only diffuse light since they have not learned this process in childhood. This suggests that temporary time reversals as analogs of the time reversed diffusion play changing A play an essential role. Note BSFRs could mean a position measurement in the space of Bohr orbits selecting a single Bohr orbit and is analogous to time reversed diffusion.

Acknowledgements: I want to than Marko Manninen and Ville-Einari Saari for inspiring discussions related to the large language models and AI in general.

REFERENCES

Mathematics

[A1] Linnainmaa S. *BIT Numerical Mathematics*, 15:165–173, 1975. Available at: `https://link.springer.com/article/10.1007/BF01932690`.

Neuroscience and Consciousness

[J1] Akyürek E et al. What Learning Algorithm Is In-Context Learning? Investigations With Linear Models, 2023. Available at: https://arxiv.org/pdf/2211.15661.pdf.

Books related to TGD

[K1] Pitkänen M. Evolution of Ideas about Hyper-finite Factors in TGD. In *Topological Geometrodynamics: Overview: Part II. https://tgdtheory.fi/tgdhtml/Btgdoverview2.* Available at: https://tgdtheory.fi/pdfpool/vNeumannnew, 2023.

[K2] Pitkänen M. Was von Neumann Right After All? In *TGD and Hyper-finite Factors. https://tgdtheory.fi/tgdhtml/BHFF.html*. Available at: https://tgdtheory.fi/pdfpool/vNeumann.pdf, 2023.

[K3] Pitkänen M. and Rastmanesh R. New Physics View about Language. In *Genes and Memes: Part II. https://tgdtheory.fi/tgdhtml/genememe2.html*. Available at: https://tgdtheory.fi/pdfpool/languageTGD.pdf, 2020.

Articles about TGD

[L1] Pitkänen M. Artificial Intelligence, Natural Intelligence, and TGD. Available at: https://tgdtheory.fi/public_html/articles/AITGD.pdf., 2017.

[L2] Pitkänen M. DMT, pineal gland, and the new view about sensory perception. Available at: https://tgdtheory.fi/public_html/articles/dmtpineal.pdf., 2017.

[L3] Pitkänen M. Some comments related to Zero Energy Ontology (ZEO). Available at: https://tgdtheory.fi/public_html/articles/zeoquestions.pdf., 2019.

[L4] Pitkänen M. A TGD based view about neuron. https://tgdtheory.fi/public_html/articles/TGDneuron.pdf., 2021.

[L5] Pitkänen M. What could 2-D minimal surfaces teach about TGD? https://tgdtheory.fi/public_html/articles/minimal.pdf., 2021.

[L6] Pitkänen M. Comparison of Orch-OR hypothesis with the TGD point of view. https://tgdtheory.fi/public_html/articles/penrose.pdf., 2022.

[L7] Pitkänen M. The possible role of spin glass phase and p-adic thermodynamics in topological quantum computation: the TGD view. https://tgdtheory.fi/public_html/articles/QCCC.pdf., 2022.

[L8] Pitkänen M. About long range electromagnetic quantum coherence in TGD Universe. https://tgdtheory.fi/public_html/articles/hem.pdf., 2023.

[L9] Pitkänen M. Are Conscious Computers Possible in TGD Universe? https://tgdtheory.fi/public_html/articles/tgdcomp.pdf., 2023.

[L10] Pitkänen M. Could neuronal system and even GPT give rise to a computer with a variable arrow of time? https://tgdtheory.fi/public_html/articles/GPT.pdf., 2023.

[L11] Pitkänen M. Deep learning from the TGD point of view. https://tgdtheory.fi/public_html/articles/TGDdeeplearn.pdf., 2023.

[L12] Pitkänen M. New result about causal diamonds from the TGD view point of view. https://tgdtheory.fi/public_html/articles/CDconformal.pdf., 2023.

[L13] Pitkänen M. About Langlands correspondence in the TGD framework. https://tgdtheory.fi/public_html/articles/Frenkel.pdf., 2024.

[L14] Pitkänen M. About some number theoretical aspects of TGD. https://tgdtheory.fi/public_html/articles/numberaspects.pdf., 2024.

[L15] Pitkänen M. Quartz crystals as a life form and ordinary computers as an interface between quartz life and ordinary life? https://tgdtheory.fi/public_html/articles/QCs.pdf., 2024.

[L16] Pitkänen M. Space-time surfaces as numbers, Turing and Goödel, and mathematical consciousness . https://tgdtheory.fi/public_html/articles/compuTGD.pdf., 2024.