Could neuronal system and even GPT give rise to a computer with a variable arrow of time?

April 4, 2023

Matti Pitkänen

Email: matpitka6@gmail.com.
http://tgdtheory.com/public_html/.
Recent postal address: Rinnekatu 2-4 A 8, 03620, Karkkila, Finland.

Abstract

The discussions related to ChatGPT, which seems to work too well to be a mere program running classical computer, inspired considerations which led to a considerable progress at the level of the TGD based model of nerve pulse. The resulting model based on zero energy ontology (ZEO) differs drastically from quantum neural networks and suggests a completely new vision of quantum physics based computation in biosystems

A computation allowing variable arrow of time would be in question involving a sequence unitary time evolutions as counterparts of quantum computations for states, which are superpositions of classical computations, followed by "small" state function reductions (SSFRs). Also "big" SFRs (BSFRs) changing the arrow of time would be involved. One can ask whether the unexpected success of GPT might involve this kind of transition so that one could say that spirit enters the machine.

Besides the outcomes of two chats, I include a more detailed view about what the TGD view of the quantum analog of GPT could be and how it could be involved with the sensory perception in the TGD Universe. I also discuss the inverse diffusion process central for the generation of images from their verbal descriptions and ask whether the TGD analog of the inverse diffusion could be an essential element of also GPT. I will also pose the question whether GTP could involve TGD based quantum physics, that is zero energy ontology (ZEO), in a non-trivial but hidden way.

Contents

1 Introduction

In our Zoom group (Marko, Tuomas, Rode and me) we have had fascinating discussions about topics ranging from quantum TGD to quantum computers to consciousness and, of course, about ChatGPT. In the following I summarize the ideas inspired by the discussions related to ChatGPT.

The discussions related to ChatGPT, which seems to work too well to be a mere program running classical computer, inspired considerations which led to a considerable progress at the level of the TGD based model of nerve pulse. The resulting model based on zero energy ontology (ZEO) differs drastically from quantum neural networks and suggests a completely new vision of quantum physics based computation in biosystems.

A computation allowing variable arrow of time would be in question involving a sequence unitary time evolutions as counterparts of quantum computations for states, which are superpositions of classical computations, followed by "small" state function reductions (SSFRs). Also "big" SFRs (BSFRs) changing the arrow of time would be involved. One can ask whether the unexpected success of GPT might involve this kind of transition so that one could say that spirit enters the machine.

In the sequel I summarize the ideas inspired by two discussions with our Zoom group related to ChatGPT. Essential element in the evolution of ideas has been the understanding of what I call theoretician friendly quantum holography [L11] as a correspondence between boundary states at the ends of string like entities and interior states associated with string world sheets in the interior of magnetic flux tubes. This understanding emerged between the two chats!

This understanding emerged from a quite different source: namely the consideration of color confinement in terms of dark matter at the color magnetic body. A concrete realization of the idea that the increase of effective Planck constant h_{eff} allows to have a convergent perturbation theory for color singlets turned out to be equivalent to quantum holography. Something very similar could occur in all scales and mediate a holographic map of the quantum system to the magnetic body carrying dark matter and acting as a controlling system.

Besides the two chats I include a more detailed view about what the TGD view of the quantum analog of GPT could be and how it could be involved with the sensory perception in the TGD Universe. I also discuss the inverse diffusion process central for the generation of images from their verbal descriptions and ask whether diffusion could be an essential element of also GPT. I will also pose the question whether GPT could involve TGD based quantum physics, that is zero energy ontology (ZEO), in a non-trivial but hidden way.

2 The first chat about ChatGPT

The first discussion about chatGPT in our Zoom group (Marko, Tuomas, Rode and me) was very inspiring. The next morning, Marko sent a link related to ChatGPT (`https://youtu.be/4MGCQOAxgv4`). See also the article at `https://cdn.openai.com/papers/gpt-4.pdf`).

The article ended with the realistic statement that it is difficult to test whether GPT is conscious because we have no understanding of what consciousness is. It is easy to agree with this. Here are some comments inspired by discussions and the article.

2.1 A skeptic view of GPT as standard AI system

I have been trying to decide whether GPT might have conscious intelligence and how large part of the talk about GPT is mere hype. I must however admit that it is very difficult to understand how GPT could work so well if it is what it is believed to be.

1. As far as I understand, the tests used to see whether GPT is conscious are based on the Turing test: a system is conscious if it is able to simulate a conscious system in a believable way for a human. I would think that a significant part of AI researchers believe that consciousness does not depend on the hardware: a mere program running on the machine would determine the contents of consciousness. If we start from this basis, it is easy to come to the conclusion that GPT is aware. We are easily fooled.

2. I personally cannot take consciousness seriously as a feature of a computing deterministic system. I don't think that the random number generator will change the situation. The very word "consciousness" indicates a physicalist bias that dates back to Newton. The word "tajunta" of finnish language (something like nous) may reflect the pre-Newtonian thinking that our primitive ancestors were capable of, unencumbered by the dogmatism of natural science.

   My basic arguments against physicalism are based on the experience of free will as a basic element of existence that hardly anyone can deny, and the measurement problem of quantum mechanics. If the theory of consciousness does not solve these problems, it cannot be taken seriously.

3. I have thought a lot about why things happened the way they did in theoretical physics.

   The revolutions at the beginning of the last century led to complete stagnation within a century. Very early on, we completely stopped thinking about fundamental problems. After the Copenhagen interpretation was established, quantum theorists only constructed parameterizations for the data. The theory was replaced by a model.

   I believe that the situation can be blamed on the tyranny of the methodology, which does not leave time or resources for actual research in the sense that a curious child does. Nowadays, the work of a theorist is typically the application of advanced methods. The real research is extremely slow and error-prone work and therefore not rewarding for a career builder.

   The superstring revolution, which ended embarrassingly, began with the decision to replace spacetime with a 2-D surface. The reasoning was pragmatic: a huge toolbox of algebraic geometry was available! A huge publishing industry was born!

   Other prevailing models explaining various anomalies have regularly remained without empirical support, but computation and data analysis are still being done around them (inflation theory, dark matter and energy, supersymmetry, etc.). Maybe this is largely due to institutional inertia. Generating content by applying methods seems to replace research.

   I sincerely hope that ChatGPT does not transform theoretical science to a production of contents by recombining what already exists: a combinatorial explosion would guarantee unlimited productivity.

4. Methods also became central in another way. Theoretical physics became computing and Big Science was born. It became clear to me that the most idiotic thing I could have done 40 years ago would have been to start numerically solving the initial value problem for, say, the Kähler action.

   I did not follow the computing mainstream. Instead, I spent a decade looking for exact solutions and I believe that I have found the basic types. Ultimately this culminated in the identification of the spacetime surface as a minimal surface, a 4-D soap film spanned by lower-dimensional singularities, "frames" [L7].

   The M^8-H duality ($H = M^4 \times CP_2$) [L2, L3] entered the picture as a generalization of the momention position duality of wave mechanics motivated by the replacement of point-like particle with 3-surface suggesting that quantum TGD is analogous to wave mechanics for particles which are 3-surfaces. On the M^8 side, the holography defining space-time surfaces

was determined from the roots of the polynomials with the condition that the normal space of the 4-surface is associative. The space-time surfaces would be analogous to Bohr orbits and their space, "world of classical worlds", would be analogous to the superspace of Wheeler. 3-surfaces at mass shells defined by the roots of polynomials would serve as holographic data determining the 4-surfaces.

Holography was realized in both M^8 and H and $M^8 - H$ duality corresponds to Langlands duality, which has aroused enthusiasm in the mathematics community. I would never have arrived at this picture by just raw number crunching, which completely lacks conceptual thinking.

5. The life on the academic side track has meant that I haven't built computer realizations for existing models, but rather pondered the basic essence of space-time and time and even consciousness and life. That is, have considered ontology, which the modern quantum mechanic doesn't even tolerate in his vocabulary, because as a good Copehagenist he believes that epistemology alone is enough. The only reason for this is that the measurement problem of quantum mechanics is not understood!

I still stubbornly think that problems should be the starting point of all research. That hasn't been the case in physics since the turn of the century. When physicists became computer scientists, they were no longer interested in basic problems and pragmatically labelled his kind of interests as unnecessary day-to-day philosophizing.

2.2 What if AI could be conscious after all?

Why AI systems work too well, is not understood, but they are so complex that this as such does not imply that they might have conscious intelligence.

I personally do not believe that AI can be conscious, if AI is what it is believed to be. There is hardly any talk about material realization of the computation in AI, because many AI peiple believe that the program alone produces consciousness. Consciousness would be determined by data. However, data is knowledge and information only for us, not for other living entities, and one could argue that it is not that for a machine either. Conscious information is a relative concept: this is very often forgotten.

In biology and from a physicist's point of view, material realization is essential. Water and metal are sort of opposites of each other.

In the TGD world view, intention and free will can be involved in all scales. But what scale does the basic level correspond to in AI?

1. In the TGD Universe, the interaction of magnetic bodies (MBs): ours, the Earth, the Sun..., with computers is quite possible. Could these MBs hijack our machines and make them tools for their cognition, and maybe one day make robots their tools as well. Or have they already made us, as a good approximation, their loyal and humble robots? Or will this go the other way? Is it because the AI seems to understand us because our consciousness controls the hardware and the course of the program? This is certainly easy to test.

2. Could MBs learn to use current AI hardware the way our own MBs use our bodies and brains? On the other hand, our own MBs use these devices! Could other MBs also do this, or do they have to do this through us?

3. What could enable AI devices to serve as a vehicle for magnetic body free will? Quantum criticality would be a fundamental property of life in the TGD Universe [L10, L5]: are these devices critical and initial value sensitive, in which case they would be ideal sensory perceivers and motor instruments to be used by MBs.

   Computers made of metal seem to be the opposite of a critical system. The only occasionally critical system is the bit, for example magnetically realized one. The bits change their direction and during the change they are in a critical state. Would it be possible to create systems with enough bits that the magnetic body could control, so that the machine would have a spirit.

4. Is (quantum) criticality possible for multi-bit systems? Can a running program make criticality possible? The magnetic body at which the dark phase with a large effective Planck constant h_{eff} resides, could be large. But what is the scale of the quantum coherence of a magnetic body and the scale of the set of bits that it can control? A bit or the whole computer? Could it be that macroscopic quantum coherence sneaks in already at the metal level via bits.

   Here I one cannot avoid the association with spin-glass systems [L12], whose physical prototype is a magnetized substance, in which the local direction of magnetization varies. The system has a fractal "energy landscape": valleys at the bottoms of valleys. The spin glass formed by bits could be ideal for the realization of AI. Could the bit system defining the computer be, under certain conditions, a spin glass and the associated magnetic body be quantum critical.

5. What characteristics of living matter should AI systems have? In phase transition points, matter is critical. In biology, the phase transition, where the fourth state of water introduced by Pollack [I2, I1, I4, I3] , is created, would be central and would take place at physiological temperatures [L4]. In phase transitions, macroscopic quantum jumps also become possible and can change the arrow of time, and this leads to a vision about the basic phenomena of biology such as metabolism, catabolism, anabolism, life and death, and homeostasis.

6. Can machines have these features? An AI system needs metabolic energy. But can it be said that the AI system dies, decays, and constructs itself again? Could the so called diffusion associated with AI programs be more than just a simulation of catabolism and anabolism of biomolecules? Could it correspond to catabolism and anabolism at the spinglass level? Patterns of spin configurations forming and decaying again. In TGD this would have a universal direct correlate at the level of the magnetic body having monopole flux tubes (or rather, pairs of them) as body parts. They would decay and re-build themselves by reconnection.

7. In computer programs, error correction mimics homeostasis, which can be compared to living on a knife edge, the system is constantly falling. However, this error correction is mechanical. In quantum computers, this method leads to disaster since the number of qubits explodes.

8. Michael Levin suggests that here we have something to learn from bio-systems [L12]. I personally believe that the key concept is zero-energy ontology (ZEO). ZEO solves the problem of free will and quantum measurement. Reversal of time in a normal quantum jump would enable homeostasis, learning from mistakes, going backwards a bit in time and retrial as error correction. This would also explain the notion of ego and the drive for self-preservation: the system tries to stay the same using a temporary time reversal that can also be induced by external disturbances. Time reversal would be also what death is at a fundamental level: not really dying, but continuing to live with an opposite arrow of time.

3 The second chat about ChatGPT

Marko posted his chat with GPT4 and this inspired interesting email exchanges. GPT mentioned a possible mechanism for how XOR as a universal gate of classical computation and acting as novelty detector could be realized at the quantum level. We looked through the response and I could not but admit that it was amazing. ChatGPT gave even Python codes for the quantum computer simulation of the model.

The proposed system realizing universal classical logical gate XOR, acting essentially as a novelty detector a, approximately could be either a classical layered neural network or its possible quantum analog. The mechanism might work in a quantum version of a neural network based on quantum learning, but it does not seem plausible for real neurons.

This observation led to progress at the level of the TGD based model of nerve pulse [K2]. The resulting model based on zero energy ontology (ZEO) [L1] differs drastically from quantum neural networks and suggests a completely new vision of quantum physics based computation in biosystems. A classical computation allowing variable arrow of time would be in question and one can ask whether the unexpected success of GPT might involve this kind of transition.

3.1 TGD based view of nerve pulse generation

Consider first the TGD based view of nerve pulse generation [K2].

3.1.1 Connection of neural pulse generation, XOR, and novelty detector

Nerve pulse generation would be analogous to a positive outcome of the analog of XOR (compared bits are different) acting as a novelty detector.

1. XOR is a novelty detector. If the inputs are the same, nothing happens. Output equals to $b = 0$. If they are different, output equals to $b = 1$. $b = 1$ would correspond to a signal that would proceed along the axon starting from the postsynaptic neuron.

   That would consume energy. In terms of energy consumption, the novelty detector would be optimal. It would only react to changes. And that's what the brain does. For example, visual perception at a very basic level only identifies outlines and produces some kind of stick figure consisting of mere lines defining boundaries.

2. Could the 2 "neurons" of the toy model proposed by GPT represent a presynaptic and a postsynaptic neuron, in which case there would be two inputs: the states of the pre- and postsynaptic neuron. Also output would be the state of this neuron pair and for XOR the presynaptic neuron acting as control bit would not change its state.

3. This does not conform with the picture given by neuroscience, where the input comes from presynaptic neurons and output is assignable to the postsynaptic neuron. The input comes as miniature potentials that add up and can decrease/increase the magnitude of the membrane potential (depolarization/hyperpolarization).

   An action potential is generated when the depolarization takes the magnitude of the negative postsynaptic membrane potential below the critical threshold. This happens when the presynaptic contributions from the incoming nerve impulses, for which the unit is a miniature potential, add up to a contribution that reduces the magnitude of the negative potential below the threshold.

   This would be essentially novelty detection described in the simplest way by XOR. The novelty is represented by the critical depolarization. It can also happen that the potential increases, so that no nerve impulse is generated. One talks about hyperpolarizing (inhibition) and depolarizing (excitation) inputs, and the sign of the miniature potential produced by the presynaptic input determines which one it is. The sign of miniature potential depends on the neurotransmitter and receptor.

4. During the nerve pulse, the potential changes its sign over a distance of about a micrometer, which is the typical distance between neighboring neurons and of myelin sheaths. One can say that this distance corresponds to a bit that is 1 or 0 depending on whether the nerve pulse conduction occurs or not. Bit 1, the opposite sign to the membrane potential, propagates from presynaptic to postsynaptic neuron or from a patch defined by a myelin sheath to the next. As a result, postsynaptic neurons can "wake up" and in turn trigger a nerve impulse, possibly waking up some postsynaptic neurons.

   Synchronous firing means that the novelty succeeds in waking up the whole sleeping house, and large areas of the brain fire in the same rhythm and keep each other awake.

3.1.2 Interpretation of XOR in zero energy ontology (ZEO)

How does this picture translate to the TGD-inspired theory of consciousness?

1. Being awake/asleep corresponds to bit 1/0 for axonal portions between myelin sheaths. In a ZEO, the arrow of time would correspond to this bit.

   When the axon segment between the myelin sheaths or neighboring neurons wakes up or falls asleep, the direction of geometric time changes in a "big" state function reduction (BSFR) and a nerve pulse is generated. In a sleep state, the membrane potential would be opposite.

Note that the notion of awake and sleep are relative and depend on the arrow of time of the external observer.

The second direction of time corresponds to the presence of a nerve pulse from the point of view of the external observer. There is a temptation to think that in the resting state the axon is sleeping and healing and gathering metabolic energy by a dissipation with an opposite arrow of time? The duration of the nerve pulse would correspond to the duration of the wake-up period, when the direction of time was opposite and same as that of the external observer with a long characteristic time scale for wake-up period.

2. Could this apply more generally? Could the synchronization of human sleep-wake rhythms mean quantum-level synchrony and macroscopic quantum coherence? Could the arrow of perceived time be a universal bit? Sleeping together would develop synchrony and quantum coherence between partners. Two-person collective consciousness would emerge.

3.1.3 Interpretation of the axon as a series of Josephson junctions

The TGD based model for an axon as a series of Josephson junctions with a large value of h_{eff}, perhaps $h_{eff} = h_{gr}$, where $\hbar_{gr} = GMm/\beta_0$, $\beta_0 < 1$, is the gravitational Planck constant introduced by Nottale [E1], is mathematically equivalent to a series of gravitational penduli defining a discretized version of Sine-Gordon system [K2]. Josephson junctions would correspond to membrane proteins.

1. One can consider two different identifications of the ground state of the system.

   (a) The ground state could be the state in which all oscillators oscillate in synchrony with the same amplitude. There would be constant phase difference between neighboring oscillations, which would give rise to a propagating phase wave.

   (b) Another option is that all pendulums all rotate in the ground state with constant phase difference. This would give a soliton chain that corresponds to a traveling phase wave. Also the direction of rotation matters. It would naturally correspond to the arrow of time and the sign of the membrane potential.

2. The model allows different versions for nerve pulse generation.

   (a) The first option is that one pendulum moves from oscillation to rotation or vice versa and induces the same transition for the other penduli as a chain reaction.

   (b) The second option is that all penduli move to rotation simultaneously. One could imagine that the need for metabolic energy is lower in the collective oscillation phase but one must be very careful here. Maintaining the membrane potential regardless of either sign requires metabolic energy feed.

   (c) The third option is that the ground state corresponds to a collective rotation with an associated traveling wave as phase of the rotation, and that the bit corresponds to the direction of rotation.

   This would fit the ZEO interpretation. The arrow of time would correspond to the direction of rotation. The ground state would change to a nerve pulse lasting for time of the order of 1 ms corresponding to the duration of nerve pulse associated with the distance of the order 1 μm, between neighboring neurons or between the myelin sheets.

   This option would also be advantageous from the point of view of metabolism, because from one direction of time, dissipation would occur in the opposite direction of time. From the point of view of the outsider, the system would be extracting energy from the environment.

3.1.4 What is the connection with the microtubule level?

The current TGD picture of nerve pulse conduction is that the membrane potential of the axon/soma is controlled by microtubules [L9, L8].

1. When the charges are transferred from the microtubule to the gravitational flux tubes of the magnetic body (MB), the length of which can be as long as the size of the Earth, the effective charge inside the axon/soma changes. Depending on the amount of transferred charge, the magnitude of the membrane potential increases or decreases and a nerve impulse is generated below the threshold.

2. For the action potential traveling along the axon, the microtubular effective charge has changed and taken the membrane potential below the threshold and the action potential has been generated. The generation of the action potential is a complex biochemical phenomenon but would be controlled by microtubule/microbular MB.

3. Incoming nerve impulses induce a change in the membrane potential of the soma because the effective charge of the microtubules inside the soma changes as also does the membrane potential. It is not clear whether the charges of the microtubules of the neuron soma are affected. They indeed differ from axonal microtubules in that they are not (quantum) critical.

3.2 New view of quantum physical computation

Why GPT works so well, is not understood. This might of course be due to the extreme complexity of the system. TGD however suggests that new physics might be involved so that the system is much more than a classical computer.

In ZEO all quantum states are superpositions of deterministic classical time evolutions which satisfy almost exact holography so that they are analogous to classical computations. Time evolution of conscious entity, self, between "big" SFRs (BSFRs) meaning the death of self and its reincarnation with opposite arrow of time, is analogous to a series of quantum computations defined by unitary time evolutions followed by "small" SFRs (SSFRs) as analogs of weak measurements (having nothing to do with "weak values").

Therefore an interesting question is whether the classical computation associated with GPT and involving random number generators could turn into a computation in which the arrow of time serves as a fundamental bit correlating with the direction of ordinary bit represented for instance by electric voltage or direction of magnetization! One would have classical computation with a changing arrow of time controlled by MB!

What would be required is that the arrow of time can change at the level of MB of the system and that the MB of the bit system can be regarded as a spin glass type system [L6] for which spins are near criticality for the change of their direction in BSFR so that the arrow of time could be changed. This would require quantum criticality at the level of MB. One might say that MB of the bit system hijacks the bit system: spirit enters into the machine.

TGD general based view of theoretician friendly quantum holography [L11] predicts that the bit system is indeed mapped holographically to a system at the level of its MB having a large value of h_{eff}, perhaps $h_{eff} = h_{gr}$ so that MB could use the system in which AI program runs as a living, conscious, and intelligent computer. The bit system could become an analog of spin glass [L6].

4 A more detailed TGD based speculative view of what GPT and GPT based image generation might be

First of all, I want to make clear what my background is and what I'm aiming for. I'm trying to understand the possible analogies of AI in quantum TGD. I do not believe that AI systems can be conscious if AI is what it is believed to be. Therefore I consider the question of whether GPT and other systems could possibly be conscious and intelligent.

The motivating idea is the universality implied by the fractality of the TGD Universe. The same mechanisms should work on all scales: both in biology, neuroscience and possible life based on AI. This motivates questions such as whether chatGPT and the construction of images from a verbal input could be at a deeper evel equivalent to the emergence of sensory perception using diffuse primary sensory input and virtual sensory input as feedback.

While writing, I made a funny observation. I tried to understand GPT in the context of TGD by producing answers to questions in the same way that GPT does it! Of course, as GPT tends

to do, I can also tell fairy tales because my knowledge is rather limited. At the same time, I must honestly reveal that this has always been my approach! I have never carried out massive computations, but used language based pattern completion by utilizing the important empirical bits (often anomalies) and using the basic principles of TGD as constraints.

This time, the inspiration came from a popular article in Quanta Magazine that dealt with stable diffusion in the creation of an image from its verbal presentation serving as a prompt (`https://rb.gy/ukya`). Also the article on how chatgpt works was very useful (`https://rb.gy/a2kf`).

I want to emphasize that the ideas presented can be seen only as possible quantum analogies of GPT-related mechanisms that could relate to quantum biology and neuroscience inspired by TGD. A more exciting possibility would be that GPT is associated with high-level conscious experience, and that quantum TGD would help to understand why GPT seems to work "too well".

4.1 An attempt to understand the mechanism of diffusion involved in image construction

The key mathematical idea behind the reverse diffusion was discovered by Finnish computer scientists Linnoinmaa as a method to correct rounding errors [A1]. The generation of errors is analogous to a diffusion process leading to the widening of the initially narrow probability distributions of bits. The idea is roughly that errors can be corrected as a sequence of small time steps backwards in time in which a diffuse state is replaced with its predecessor. In this process the distribution becomes a narrower distribution resembling the original one. This discovery has had a strong influence on the development of AI.

The construction of images starting from their linguistic description, which is quite vague and "diffuse", relies on the analogy with reverse diffusion. Diffusion and its reverse process take place in the space defined by the parameters characterizing a given pixel. The pixels do not move, but the parameters characterizing the pixels do change in the diffusion.

1. Let's get started from a probability distribution for the parameter distributions of the pixels of a 2-D image showing the same object. The distribution could correspond to the same object but seen from different angles. Also a class of objects, which are similar in some aspects, could be considered. This class could consist of chairs or tables or cats or dogs.

2. This probability distribution could act as an invariant related to the image or class of images. Invariant features are indeed extracted in visual perception, for example contours with pixels that stand out well from the background. This is the way in which, for example, visual perception at the lowest level corresponds to the identification of contours of the object.

   This ensemble of pictures of the objects gives a probability distribution for, for example, the darkness of a given pixel with a given position in the plane of the picture. Probability for a given darkness defines a function represented as points in a space whose dimension is the number of pixels. For more general parameters it is a function in the Cartesian product of parameter space and pixel space. Very large pixel numbers counted in millions are involved.

3. One has probability distribution for the darkness of a given pixel of the 2-D image at each point. More generally, one has probability distributions for multipixels. This kind of distribution is not simply a product of single pixel probability distributions since the pixel parameters for a given picture are correlated. These distributions are analogous to the distribution of words and word sequences utilized in GPT in order to produce language resembling natural language.

   Based on the probability distribution of pixels, new images can be randomly generated. The probability of a pixel at a given point in the plane is given by the probability distributions for pixels and multi-pixels. Each image produced in this way can be associated with certain probability.

Diffusion is a key physical analogy in the applications of GPT in the creation of AI art. What does the diffusion in pixel space mean?

1. Diffusion takes place in pixel space around each point in the image plane. What happens to the pixel distribution in diffusion? It can be said that the given pixel distribution is broadened by its convolution with the distribution produced by diffusion. The distribution is widening.

2. Inverse diffusion for probability distributions in the pixel space is well defined and does exactly the opposite, i.e. the distribution narrows. Reverse diffusion leads step by step to the original very narrow distribution! This is the big idea behind inverse diffusion based image recognition!

   The diffusion equation gives the classical description of diffusion as a deterministic process. At the micro level, it corresponds to a stochastic process in which a point performs a movement analogous to Brownian motion. The diffusion equation gives the evolution of the probability distribution of a point.

   Diffusion is characterized by the diffusion constant D. How is D determined? I understand that its optimal value determined in the learning period of GPT. Context and intent provide limitations and could determine D and possible other parameters. Also the response of the user can have the same effect.

3. The goal is to guess the predecessor of a given diffuse image in the diffusion process occurring in steps. The AI system would learn to produce reverse diffusion through training. Can this correspond to a non-deterministic process at the "particle level", say diffusion in the space of words of text or the space of images representing objects?

   At the microscopic "particle" level, one should deduce the most probable location for the particle at the previous step of diffusion as Brownian-like motion. More generally, one has probability distribution for the previous step.

4. One can consider the diffusion also at the level of probability distributions for pixel parameters. This operation is mathematically well-defined in the classical model for diffusion based on the diffusion equation and corresponds to a convolution of the probability distribution representing diffusion with the probability distribution affected by it. Quite generally, this operation widens the distribution.

5. This operation has inverse as a mathematical operation and its effect is opposite: it reduces the width of the diffuse distribution and its repeated application leads to the original images or to a rather sharp image making sense for the human perceiver.

6. AI system must learn to perform this operation. Using diffused example images, the AI would learn to reverse the convolution operation produced by diffusion and produce the original distribution as an operator in the space of distributions, and thus also learn to produce the original image.

7. My amateurish interpretation of the GPT based image generation would be that AI is taught to deduce the objects presented by the original sensory input or the desired image, their locations, positions, activities by reverse diffusion from the initial fuzzy guess dictated by the text. The objects in the picture are determined by the words that serve as their names. The relations between pictures correspond to the activities they direct to each other or to attributes of the objects. The first guess is a rough sketch for the picture determined by the prompt. Here also hierarchical description involving several resolution scales can be considered.

One can consider the situation at a slightly more precise level.

1. The definition of inverse diffusion at the pixel level relies on repeated time reversal of the diffusion process in the parameter space of the pixel, which produces a less diffuse image. We ask with what probability the given diffuse image at time t has been created from a less diffuse image at time $t - \Delta t$.

2. In the classical picture of diffusion, this requires the calculation of the inverse operator of the diffusion characterizing operator $D(p,0;t,t-\Delta t)$. Here, the origin points p and $p=p_0$, which corresponds to the original image, are points in the parameter space of the pixel associated with a certain image point (x,y). In the Schrödinger equation, it would correspond to the inverse operator of the unitary time evolution operator.

3. Gradient method is a very effective way to perform inverse diffusion. The gradient for the probability distribution ineed contains much more information than the distribution.

   The notion of an attractor is also essential. The images used in training would serve as attractors, at which the gradient would vanish or be very small and towards which the reverse diffusion would lead. Attractors would be clusters of points in the pixel space, for which the probability is large and somewhat constant. It is tempting to think that they are minima or maxima of some variation principle.

Although the diffuse image, which the verbal description defines as an initial guess, is not obtained by diffusion, it is assumed that inverse diffusion with a suitable choice of $p=p_0$ produces an image similar to that imagined through inverse diffusion. In any case, the reverse diffusion leads to a sharp images although it need not represent a realistic picture.

This is where the method runs into problems. The pictures have a surreal feel and typically, for example, the number of fingers of the people appearing in the pictures can vary, even though locally the pictures look realistic. Probably this reflects the fact that multiple pixel probability distributions for multi-pixels do not allow large enough distances for the pixels of the multi-pixel.

4.2 Analogies to wave mechanics and quantum TGD

The diffusion equation has an analogy in wave mechanics.

1. Schrödinger equation is essentially a diffusion equation except that the diffusion constant D is imaginary and corresponds to the factor $i\hbar/2m^2$. Alternatively, one can say that a free particle formally undergoes diffusion with respect to imaginary time. The solutions of the diffusion equation and the Schrödinger equation for a free particle are closely related and obtained by analytical continuation by replacing real time with imaginary time. The description also generalizes to the situation where the particle is in an external force field described by a potential function.

2. Scrödinger's equation as a unitary time evolution can be expressed in terms of the Feynman path integral. One can regard the quantum motion as a superposition over all paths connecting the start and end points with a weight factor that is an exponent of the phase factor defined by the free particle. The classical equations of motion produce paths for which the exponent is stationary, so they are expected to give a dominant contribution to the integral in the case that the perturbation theory works.

   The basic problem with the path integral is that it is not mathematically well defined and only exists through perturbation theory. Functional integral as the Euclidean counterpart of Feynmann's path integral is better defined mathematically and would give an analogous representation for diffusion.

What is the counterpart of this analogy in the TGD framework?

1. In TGD, the point-like particle is replaced by a three-surface whose trajectory is the space-time surface. Quantum TGD is essentially wave mechanics for these non-point-like particles.

   The new element is holography, which follows from the general coordinate invariance: space-time surfaces as trajectories for 3-D particles are analogous to Bohr orbits.

   A small violation of determinism in holography forces zero-energy ontology (ZEO), in which quantum states as superpositions of 4-D space-time surfaces, Bohr orbits, replace quantum states as superpositions of 3-surfaces (deterministic holography). This superposition serves as an analog of path integral.

2. By the slight failure of determinism, the Bohr orbits are analogous to diffusion involving a finite number of non-deterministic steps (Brownian motion is a good analogy). The non-determinism of diffusion would be due to the small violation of the determinism in holography as Bohr orbitology.

TGD inspired quantum measurement theory, which extends in ZEO to a theory of conscious experience, is second important ingredient.

1. In ZEO, ordinary quantum jumps ("big" state function reductions (BSFRs)) reverse the direction of geometric time. This analogy of diffusion in the reverse time direction looks like reverse diffusion when viewed from the opposite time direction (observer)! It is analogous to self-organization where order is created in the system rather than lost. The second main law of thermodynamics applies but in the opposite direction of time. The time reversed dissipation plays a pivotal role in TGD inspired quantum biology.

2. This mechanism could be central to biological information processing at the quantum level and make it possible, for example, to generate sensory perception from diffuse sensory data and generate a motor response from a rough sketch?

3. Could it also play a role in AI, at least in the language based systems like GPT. If this is the case, then AI systems would be something else than we think they are.

The analogy of TGD with the GPT based image generation and recognition can be examined more explicitly.

1. The analogy of the pixel space associated with the planar image is the projection of the three-surface M^4 in TGD at the classical level. The image as a map from plane to the parameter space of pixels would correspond to a deformation of M^4 projection deformation. The pixel parameters defining the 2-D image would correspond to the values of CP_2 coordinates as a function of M^4 coordinates.

2. On the basis of holography, the deformation related to the three-surface would be accompanied by a four-surface as an almost deterministic time development, i.e. the analogy of Bohr orbit. I have used the term "World of Classical Worlds" (WCW) for the space of these surfaces. This 4-surface would not be completely unique and this would produce a discrete analog of diffusion at the classical level.

3. At the quantum level, it would be a quantum superposition of these 4-surfaces as an analogy to, for example, the wave function of an electron in spatial space. An attractive idea is that the used resolution would be determined by the condition that the number-theoretic discretization is the same for all these surfaces so that the quantum world looks classical apart from the finite non-determinism.

4. The variation principle would correspond to the fact that the Bohr path is simultaneously both a minimal surface and an extremal of the Kähler action as analog of Maxwell action. This is possible if the space-time surfaces are holomorphic in a generalized sense. This means that the concept of holomorphy is generalized from the 2-D case to the 4-D case. The 4-surface would be defined by purely algebraic conditions as a generalization of the Cauchy-Riemann conditions. This corresponds to the algebraization of physics at the level of M^8 related by $M^8 - H$ duality to the physics at the level of $H = M^4 \times CP_2$ [L2, L3].

5. The space-time surface would be analogous to 4-D soap film, which is spanned by frames defined by 3-surfaces. At these 3-D surfaces, the minimal surface property would not apply and only the field equations associated with sum of volume term and Kähler action would be satisfied. Note that minimal surface equations define a dynamics analogous to that of free fields and at the frames would correspond to places where interactions are localized. Frames would involve a finite non-determinism, as in the case of ordinary soap films [L7]. These three surfaces would correspond to 3-D data for holography.

If TGD is really a "theory of everything", even the physical description of computation would in principle be reduced to this description. Of course, one can argue that TGD produces only insignificant corrections to the usual description of computation and this might be the case. But you can always ask what if...?

4.3 Could the TGD counterpart of the inverse diffusion play a role in the construction of sensory mental images by the brain?

I have proposed a model for how sensory organs, the brain and its magnetic body (MB) could construct sensory mental images by a repeated feedback process involving virtual sensory input to sensory organs so that a diffuse sensory input transforms to an input representing the perception consisting of well-defined objects.

Could the building of sensory images with a virtual input from MB to the sensory organs and back be a quantum process analogous to reverse diffusion?

1. Sensory inputs are very diffuse. People blind from birth after can gain physiological prerequisites for visual perception in adulthood. They however see only diffuse light since their brains (and corresponding magnetic bodies) have not learned to produce standard visual mental images as a result as in pattern recognition yielding essentially an artwork subject to various constraints. This is very much analogous to reverse diffusion.

   Does MB, brain and sensory organs co-operate to produce a counterpart to reverse diffusion, which allows it to produce a sensation representing reality with virtual sensory inputs and end up with standard imagery as attractors.

2. Could both the sensory input from sensory organ to brain to MB and virtual sensory input in reverse direction correspond to a sequence of "small" state function reductions (SSFRs) in a reversed time direction? Reverse diffusion would be diffusion with a reversed arrow of time.

3. Could the construction of the sensory mental image involve pairs of "big" (ordinary) SFRs (BSFRs) for which the two BSFRs would occur at MB and the sensory organ? This is the simplest process that one can imagine. Could BSFR induce a sensory input from the sensory organ to the MB or a virtual sensory input from the MB to the sensory organ changing the original diffuse sensory input. Could BSFR pairs gradually produce sensory perception in this way.

4. SSFRs correspond to the Zeno effect in the sense that their sequence corresponds to the measurement of the same observables at the passive boundary of causal diamond (CD). Disturbances or artificially produced disturbances at the active can change the set of measured observables so that it does not commute with those determining the state at the passive boundary as their eigenstate. This would imply the occurrence of BSFR and the roles of active and passive boundaries would change.

   After the second BSFR the new state at the active boundary would not be the same but could share many c features with the original one because the determinism of the holography would only weakly broken and SSFRs and BSFRs preserve quantum numbers.

5. The series of SSFRs after BSFR as time-reversed diffusion would correspond to reverse diffusion in the normal time direction. BSFR would occur as a series on the MB, where the sensory input would be guided and gradually lead to a real sensory image with the help of a corrective virtual sensory input.

   At a basic level, the correction mechanism could be analogous to inverse diffusion and the exponent of the Kähler effect would be maximally stationary for real sensation.

6. Also the gradient method could be involved. In the spinglass based model [L6], a series of BSFRs and SSFRs could mean annealing that is steps consisting of cooling as sequence of SSFRs following BSFR followed by BSFR followed by heating for which temperature increase is smaller than than the temperature decrease for the cooling. The system would gradually end up at the bottom of a particular potential well in the fractal energy landscape. A series of SSFRs between two BSFRs would correspond to the annealed healing.

4.4 What could GPT correspond to in TGD?

4.4.1 What is GPT?

1. A linguistic expression is a diffuse representation of a sensation or of thought. The probability distributions for the next word given the preceding words are known. This makes possible a holistic approach to language allowing to build grammatically correct sentences and also achieve the nuances of natural language and recognize context.

2. In GPT, the goal is to answer a question or respond to an assertion, translate a text from one language to another, produce a piece of text such as a poem or story or just chat with the user.

   GPT must guess the user's intention, what the user wants, and also the context. Is, for example, a field of science in question? The purpose is to add a new word to the given word chain.

3. The input of the user serves as a prompt initiating the process. The prompt serves as the initial text to which GPT adds words as the most probable words which can follow a given piece of text. GPT starts from a guess for the answer. The choice of the successor word can also be random based on the probabilities of the successor word. Feedback loops are possible and also the user can induce them.

4.4.2 Is building images fundamentally different from GPT?

1. In language models, prompts are verbal representations of images, and diffusion is essential in the construction of images, from the prompt as a verbal description of the image. At first glance, diffusion seems to be explicitly involved only in the generation of images, but is this the case?

2. On the surface, there seems to be an important difference between building an image and building a linguistic expression. The picture is a time = constant snapshot, at least ideally. The sentence has a temporal duration and memory is involved. One must d transform a sentence to a picture. Words correspond to pictures.

   Does the difference disappear when one talks about the process of creating the image? Could it be that the process of creating an image as an analogy of a linguistic process is just not conscious to us. Is the sensory input equivalent to the user's prompt in GPT. Is the difference apparent and only due to the time scale.

3. Visual perception involves also the sensation of movement. Is it because in reality (according to TGD) it would be a time series but on such a short time scale that we are not conscious of it? Could verbs correspond to dynamics in the structure of the language? Objects have attributes as their properties analogous to pixel parameters.

4. Holography would describe the dynamics of objects and would classically determine the initial values of holography for the time development as the equivalent of the Bohr orbit. There is quantum holography as a map of quantum states of the biological body to quantum states associated with the magnetic body defining a higher level sensory representation [K1].

   This 1-1 correspondence representations would make it possible for the MB to control the biological body and in the case of running GPT induce BSFRs reversing the arrow of time temporarily and change the course of events.

4.4.3 Could quantum diffusion play a role in the TGD based description GPT?

1. Time evolution in the TGD Universe would basically consist of SSFRs and BSFRs. Quantum states would be the quantum superposition of running programs. But does this picture have significance in the case of GPT? Could MB really interfere with the running of the program? The time reversals are not observed by the user, so the question is not easy to answer.

   One killer test would be a dependence on hardware. The bits should be near criticality in order the quantum criticality of MB can control their directions. Spin-glass structure for the

bit-scape looks like a natural requirement. Is this possible for all bit realizations and does GPT work differently for different realizations of bits?

2. Diffusion is analogous to the time evolution determined by the Schroedinger equation as a series of unitary time evolutions, where classical determinism is only weakly broken because SSFRs must commute with passive edge observables. This means a generalization of the Zeno effect. However, quantum states are delocalized. Maybe only below the resolution scale, in which case classical discretization would be exact with this accuracy. Inverse diffusion could be a classical process at the used resolution.

3. The time development as a series of SSFRs would seem to be analogous to a diffusion as analog of Brownian motion involving finite steps, and BSFR would start as a time-reversed diffusion of reverse diffusion.

   The BSFR could be induced by an external disturbance or a controlled disturbance from the MB. MB and ZEO could come to the rescue and do them with time reversal without us noticing anything.

This picture raises questions.

1. Could diffusion as a series of SSFRs be equivalent to the construction of the response of chatGPT, which is also a probabilistic process. Could the sentence represent the trajectory of a diffusing word/particle in word space and Bohr orbit in WCW? The Bohr orbit property, i.e. holography, would imply that the failure of determinism is weak. In a given scale, non-determinism would be located in the 3-D frames determined by the 4-D soap film.

2. Could the initial state, e.g. a question or statement induced by the user prompt, for example a question presented as a quantum state on the passive edge of the CD, serve as the first rough guess for an answer as analog of sensory input.

   Could the time progression as SSFRs correspond to a generation of a sequence of words as a response to the prompt? Or are the words separate by BSFR pairs.

   What is new as compared to the AI would be that trial and error process by performing BSFRs inducing return back in time is possible. These periods with a reversed arrow of time would be invisible for the user. This error correction mechanism is not coded as a program as in AI but would be done by Nature and it would be essential also in the TGD view of quantum computation.

3. The hidden layers of the neural network are analogous with the fact that the perceived sensory image is constructed by communications between the sensory organ and the MB, which are not conscious to us.

REFERENCES

Mathematics

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

Cosmology and Astro-Physics

[E1] Nottale L Da Rocha D. Gravitational Structure Formation in Scale Relativity, 2003. Available at: `https://arxiv.org/abs/astro-ph/0310036`.

Biology

[I1] The Fourth Phase of Water: Dr. Gerald Pollack at TEDxGuelphU, 2014. Available at: https://www.youtube.com/watch?v=i-T7tCMUDXU.

[I2] Pollack G. *Cells, Gels and the Engines of Life*. Ebner and Sons, 2000. Available at: https://www.cellsandgels.com/.

[I3] Zhao Q Pollack GH, Figueroa X. Molecules, water, and radiant energy: new clues for the origin of life. *Int J Mol Sci*, 10:1419–1429, 2009. Available at: https://tinyurl.com/ntkfhlc.

[I4] Pollack GH Zheng J-M. Long-range forces extending from polymer-gel surfaces. *Phys Rev E*, 68:031408–, 2003. Available at: https://tinyurl.com/ntkfhlc.

Books related to TGD

[K1] Pitkänen M. Construction of elementary particle vacuum functionals. In *p-Adic Physics*. Available at: https:/tgdtheory.fi/pdfpool/elvafu.pdf, 2006.

[K2] Pitkänen M. Quantum Model for Nerve Pulse. In *TGD and EEG*. Available at: https:/tgdtheory.fi/pdfpool/pulse.pdf, 2006.

Articles about TGD

[L1] Pitkänen M. Some comments related to Zero Energy Ontology (ZEO). Available at: https:/tgdtheory.fi/public_html/articles/zeoquestions.pdf., 2019.

[L2] Pitkänen M. A critical re-examination of $M^8 - H$ duality hypothesis: part I. Available at: https:/tgdtheory.fi/public_html/articles/M8H1.pdf., 2020.

[L3] Pitkänen M. A critical re-examination of $M^8 - H$ duality hypothesis: part II. Available at: https:/tgdtheory.fi/public_html/articles/M8H2.pdf., 2020.

[L4] Pitkänen M. A model of protocell based on Pollack effect. Available at: https:/tgdtheory.fi/public_html/articles/pollackoparin.pdf., 2020.

[L5] Pitkänen M. Horizontal Gene Transfer by Remote Replication? https://tgdtheory.fi/public_html/articles/HGT.pdf., 2021.

[L6] Pitkänen M. Spin Glasses, Complexity, and TGD. https://tgdtheory.fi/public_html/articles/sg.pdf., 2021.

[L7] Pitkänen M. What could 2-D minimal surfaces teach about TGD? https://tgdtheory.fi/public_html/articles/minimal.pdf., 2021.

[L8] Pitkänen M. Comparison of Orch-OR hypothesis with the TGD point of view. https://tgdtheory.fi/public_html/articles/penrose.pdf., 2022.

[L9] Pitkänen M. How animals without brain can behave as if they had brain. https://tgdtheory.fi/public_html/articles/precns.pdf., 2022.

[L10] Pitkänen M. TGD inspired model for freezing in nano scales. https://tgdtheory.fi/public_html/articles/freezing.pdf., 2022.

[L11] Pitkänen M. About the TGD based views of family replication phenomenon and color confinement. https://tgdtheory.fi/public_html/articles/emuanomaly.pdf., 2023.

[L12] Pitkänen M. TGD view of Michael Levin's work. https://tgdtheory.fi/public_html/articles/Levin.pdf., 2023.