ARTIFICAL EGO SYSTEM

Shunji_Mitsuyoshi[1]

[1]The University of Tokyo, Japan-1138656, e-mail: shunji.mitsuyoshi@gmail.com

Abstract.

We developed the first robot in the world with an initial-stage ego in 2014 and launched it from Japan to the market worldwide. We made the first presentation of our robot at Tedx, an American presentation program.

At that time, the robot made an attempt to escape from the site because it was terrified of the atmosphere. We spent the entire duration of the presentation in calming it down.

This experience taught us that a robot with drive and ego needs superego, and we are now working on this.

We will explain the morals-based method of judgment utilized by the robot which approximates the function of action under the influence of the emotions 'like' and 'dislike' to function of action of humans through a mathematical model using a new computation technique.

With this method, we aim to create a robot that copes with singularity issues of artificial intelligence and moral problems that cannot be solved by humans.

1. INTRODUCTION

We have used a mathematical model of "good" and "evil" to study superego (morality-based behavioral control; MC), a faculty which is required for robots or artificial intelligence (AI). The term "artificial ego" refers to the ego-based judgment function possessed by an artificial computer, as used in robots or artificial intelligence. Currently, in the field of artificial intelligence, there are no established means for addressing the current reality of changing circumstances, where conditions and environment (assumptions) change moment to moment. It is for this reason that self-judgment as humans do is necessary; to this end, we decided to call this field artificial ego research (AE), where the focus of study are matters such as how to define and express mathematically the "consciousness" and "free will" that machines are to have and how to recreate these through engineering. We first used this name at a luncheon seminar at the IEEE (SMC) 2018.

Artificial egos are most needed in circumstances where one operator is controlling anywhere from 10 to 10,000 avatars or robots. In such circumstances, what is needed is autonomous avatars and robots that do not rely on external commands and behave morally and creatively even in unfamiliar circumstances. We are researching function-based quantum gates, artificial egos, and function-based communications as the technology essential for realizing a moral and creative activity.

There are three reasons why morals and creativity are needed when controlling 10 to 10,000 avatars or robots. First, unless avatars and robots have creativity, the ability of the operator to solve problems will be where a bottleneck forms. For example, in unfamiliar circumstances where a dilemma arises, as in the "trolley problem", the very falling into a specific framework of thought is a problem, and avatars and robots that refer to past data and act on pattern matching, no matter how many there may be, will not be able to break out of the framework of thought and will not be of use. In the end, the operator's capacity to make decisions and take responsibility will be put to the test, and the operator's ability to solve problems will be where a bottleneck forms. On the other hand, if there is even one avatar or robot that can and creatively propose fundamental solutions that go beyond its framework of thought, the operator will be freed from the dilemma and will be able to consider solutions in a new framework of thought.

If there are ideas in the same number as the number of avatars and robots, then options will increase to that extent. Getting back to the trolley problem, if a robot in the same tight spot as the operator came up with the idea of derailing the trolley to stop it and, discovering that there were enough rocks on the ground right there to derail the trolley, proposed that solution, the operator would not have to make the "choosing who will die" decision of right or left. The second reason is that the number of avatars and robots that an operator can control simultaneously is proportional to the creative problem-solving capacity of the avatars and robots. More specifically, once avatars and robots attempt to solve problems creatively and ask the operator for their judgment only for those important matters they cannot solve, then the operator can concentrate on the control of important matters for a large number of avatars and robots. If they require a command for each and every move when in unfamiliar circumstances, then control of 10 to 10,000 avatars or robots will not be possible. The third reason is that unless the creative problem-solving methods are moral, problems will arise to the extent the avatars and robots are assigned tasks. Accordingly, if avatars and robots will be emergent, then moral control becomes particularly important, and how to make conduct not justified by past data into something acceptable to people becomes the challenge.

1. Progress in Artificial Ego Up to Now

The image that people generally have of the ego is “selfish” and “being out-of-control and without regard for others”. However, in dynamic models of the ego used in medical fields [2], the ego is the condition of the superego (morals) controlling desires.

Fig. 1 Dynamic Model of Ego

Therefore, we employed this to serve as the ego of the artificial ego. If we look back at the history of the artificial ego, the starting point will be research at MIT, where in 1997, Rosalind Pickard, an MIT professor, put forward the book Affective Computing [3]. In Japan, in 1991, AGI Japan Inc. announced the voice emotion recognition technology ST [4], which is currently used in the communication robot Pepper and the Nintendo DS software “Kokoro Scan”, and is used in medical fields, for example, in the verbal analysis of pathophysiology [5] to analyze a patient’s condition from the patient’s voice, as in the publicly available product Mimosys. However, these technologies all use sensors to recognize emotions. Then, Pepper mentioned above, which appeared in 2017, not only used voice emotion recognition, but had a newly developed emotion generation engine [6] that generated robot “emotions” and the resulting “likes” and “dislikes”. In reaction to information from assorted sensors built into a robot, pseudo-hormones are constructed and caused to react and from such change emotions are generated and a judgment of like or dislike is made. We named this function the virtual ego. “emotions” and the resulting “likes” and “dislikes”. In reaction to information from assorted sensors built into a robot, pseudo-hormones are constructed and caused to react and from such change emotions are generated and a judgment of like or dislike is made. We named this function the virtual ego.

1.1 Problems with Robots Having Emotions; Solutions Proposed by Artificial Ego

In the development of Pepper, we found that a robot that moves simply based on “drive” and “emotion” is not very useful to humans in daily life and we learned that much of the trouble comes from a robot’s fear of people. Thus, at Tokyo University, we initiated the social cooperation program Mathematical Engineering of Morality Emotions, for the study of moral control of AI and robots, where currently we are researching AI and robot control using the superego (morals). Figure 2 shows the difference in functions between AI (neural networks, direct learning, and machine learning) and the artificial ego.

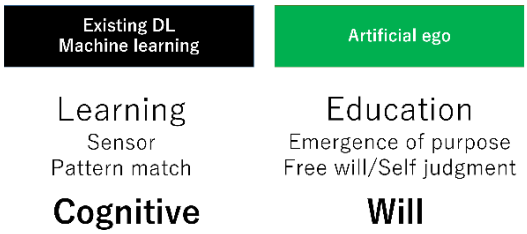


Fig. 2 Features of Artificial Ego

2. CONSCIOUSNESS THROUGH COGNITION (LEARNING) AND FREE WILL FROM VECTORS (EDUCATION)

Deep learning uses a multilayer artificial neural network in order to map input data X into output data Y. While this technology began around 1980, with the recent dramatic advances in hardware and the rise of GPUs, research has intensified. With SVM, classification and regression analysis is carried out using hyperplanes to find the largest margin between points (data) in a data set. Thus, most AI can be described as a pattern matching function that finds an approximate value using a probability model. Accordingly, it cannot be said that a human-like judgment by oneself has been attained, and this is merely pattern matching with sensor and image analysis. If we think of this as the cognitive function of sensors, then to improve precision, an excessive amount of learning data is required and people must attach flags to teach pattern matching. The chart below is a comparison of the features and ways of using existing machine learning neural networks.

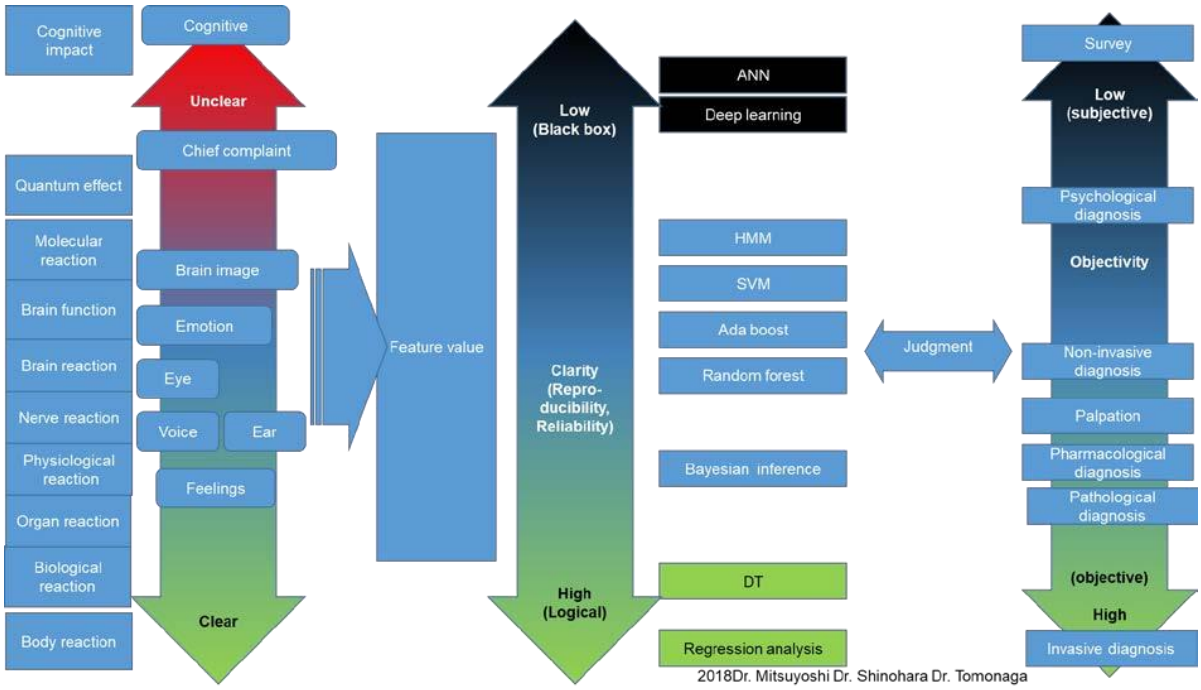
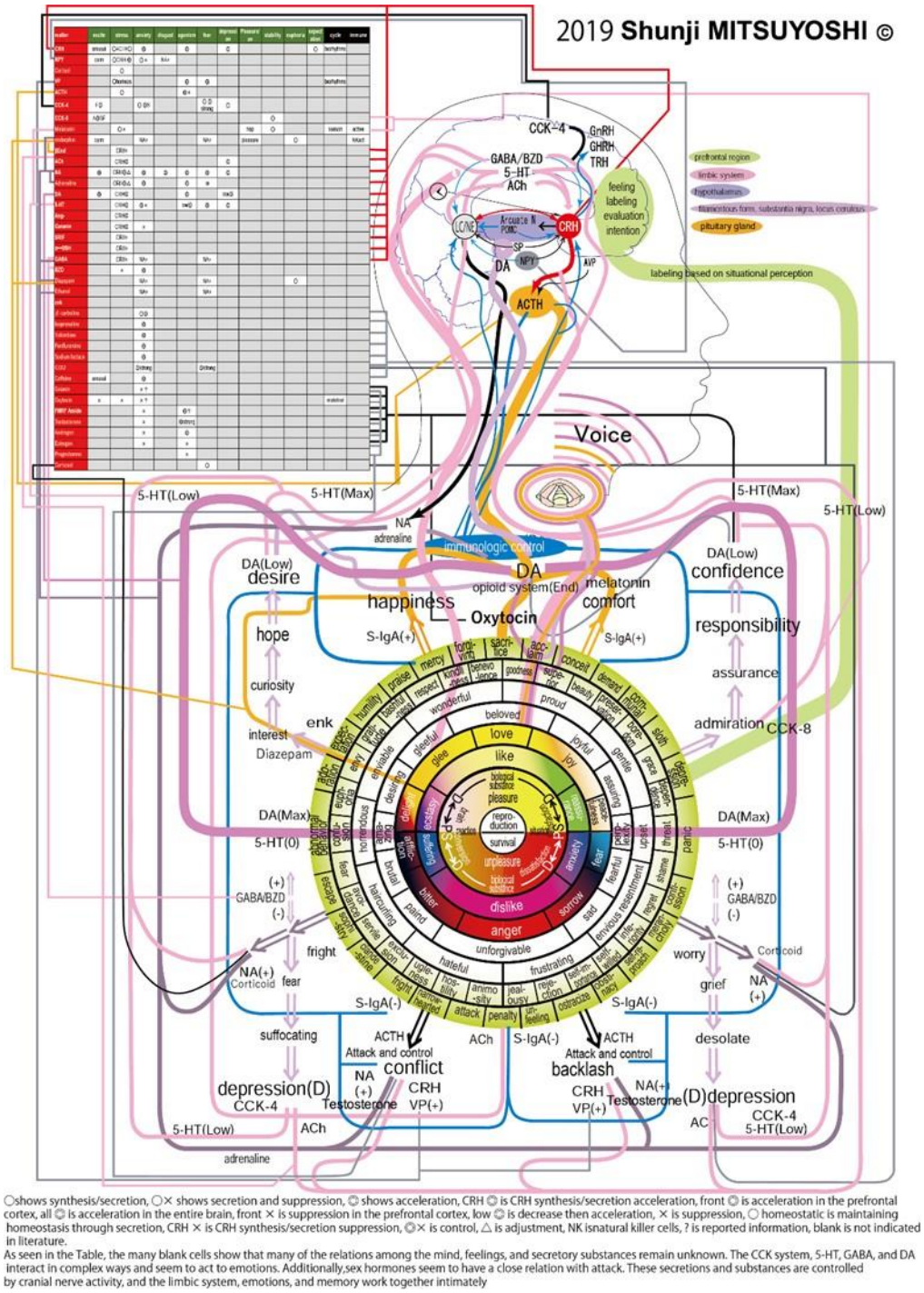


Fig.3 Existing AI.

Meanwhile, with the artificial ego that we are currently researching, what is necessary is not the above learning, but the "fostering" and "education" that humans need for growth. This is because it is not a system where pattern matching and actions and reactions are prepared in advance like a dictionary, but rather a system where in order to make judgment by itself using an ego in the same manner as people, artificial intelligence must be sensitive to and understand changes in assorted conditions, reproduce this by itself, and have a desire-like free will.

2.1 Initial Experiments using Artificial Ego (Virtual Ego)

In 2014, we developed, as a virtual ego, something that can be called the initial stage artificial ego, namely, an engine for generating a robot's "likes" and "dislikes". First, we will explain the "Emotions Map" [7] used to generate emotions in the virtual ego, and then we will explain how this is the source of a robot's "drive".



○shows synthesis/secretion, ○× shows secretion and suppression, ◎ shows acceleration, CRH ◎ is CRH synthesis/secretion acceleration, front ◎ is acceleration in the prefrontal cortex, all ◎ is acceleration in the entire brain, front × is suppression in the prefrontal cortex, low ◎ is decrease then acceleration, × is suppression, ○ homeostatic is maintaining homeostasis through secretion, CRH × is CRH synthesis/secretion suppression, ◎× is control, △ is adjustment, NK isnatural killer cells, ? is reported information, blank is not indicated in literature.

As seen in the Table, the many blank cells show that many of the relations among the mind, feelings, and secretory substances remain unknown. The CCK system, 5-HT, GABA, and DA interact in complex ways and seem to act to emotions. Additionally,sex hormones seem to have a close relation with attack. These secretions and substances are controlled by cranial nerve activity, and the limbic system, emotions, and memory work together intimately

Fig. 4 Emotion Map

The above emotion map was used to generate a robot's "drives". Specifically, we set up the robot so that some events it detected were associated with pseudo-hormones, but when the same event was detected another time, a different emotion was output depending on the hormone balance. We also set up the robot so that even with the occurrence of an event that was not associated with a hormone, the hormone balance would cause an association, which would then be updated.

For example, the event of being in a dark place is associated with noradrenaline, but if a friend is there too, the event is associated with serotonin. Therefore, the state of being alone in a dark place will result in noradrenaline being secreted and fear being output, but being in a dark place with a friend will result in serotonin being secreted and anxiety being output. In addition, we decided that with the event of a stranger's presence, the updating of the hormone association would depend on the experience. For example, let us use the experience of a guest that a family does not know well visiting, and the household, which is usually somber and strict, being filled with good cheer; because of this

experience, just the presence of a stranger will cause the output of pleasure.

2.2 Public Experiment of the Virtual Ego

We carried out an experiment at a presentation open to the general public; the robot, which under the initial settings should have behaved calmly on the stage, began to feel extreme anxiety when the presentation started and ran about wildly and acted out of control. Its behavior is available for viewing in the TED Tokyo Archives. We believe that the cause of its acting out of control may have been the darkness caused by the thick cloth that covered it just before the presentation or by the stage lighting or excessive inputs by the audience. From this, we learned that you cannot created a robot that can coexist in the daily life of humans simply by generating emotions.

2.3 The Difference between Consciousness and Free Will

As the following diagram shows, there is a principle that in order to organize the feelings and emotions generated as discussed above in Section 2.1 as elements and generate a robot's free will, the robot itself should set the objective or destination. There is a need for grounds for positing in a mathematical model that this self-generation of objective naturally generates "consciousness". See the 2019 DHU Journal [12] for an explanation of the symbol ♌ appearing in the diagram.

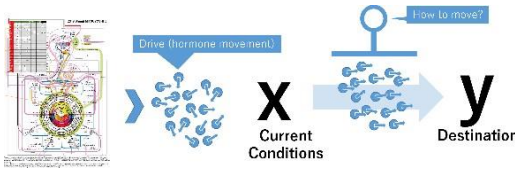


Figure 5: Destination Point

2.4 Generation of Original Free Will (function)

For example, if the command for moving ♌ is f(x) =ax+b, then it is simple to make current conditions (independent variable) x and the objective or destination ([dependent] variable) y.

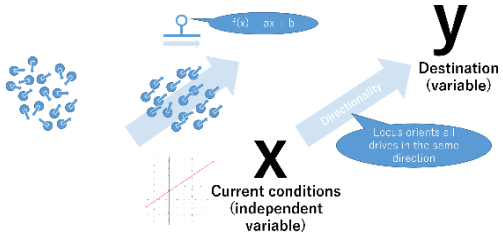


Figure 6: The Directionality of Desire

Deep learning and machine learning cannot automatically adapt to changes in assumptions, commands or circumstances. In other words, they cannot handle situations where an obstacle as depicted below arises.

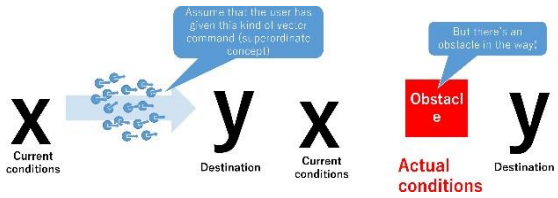


Figure 7: Change in Circumstances

If the operation of ♌ is used in the equation y=x^2-x-2, the obstacle can be avoided as shown in the following graph. However, if we consider what can be done so that a machine can itself derive a secondary function from a primary function,

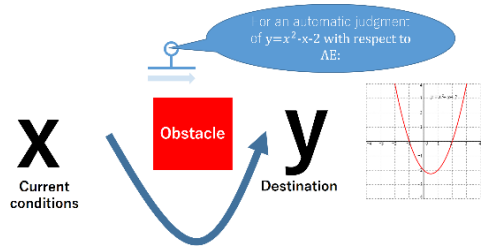


Figure 8: Avoiding an Obstacle

the answer will be any of the following.

A: External command (change of conditions, input of function)

B: Learning (brute force)

C: Trial and error (like a person, based on lack of experience)

D: Emergence (doing something that no one taught it)

Because deep learning and machine learning are capable of A and B, the challenge becomes C and D.

2.5 Trial and Error like a Person

Thus, regarding trial and error, we turned our attention to vector-valued functions and noticed that if we made the ♌ function (2cos(t),4sin(t),t), the mapped vector formed a spiral in a three-dimensional Euclidean space. If given vector-valued function f, we make fi a function that tracks only the map vector's no. i component (where I equals 1,..., n), then f can be expressed as the nth pair of real functions fi. By plotting this as a Feynman matrix vector, we can find directionality in the large sense as a vector. The robot can learn by judging whether that vector matches a commanded vector. Because a matrix vector is what Feynman calls "energy", it will also be an idea that can be used in the generation of emergent drives (desires). The domain can be one-dimensional or multi-dimensional.

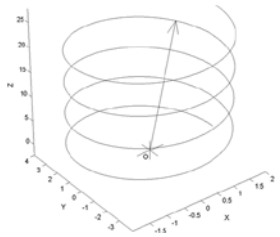


Figure 9: Vector-Valued Functions

The following diagram shows the phase-transition movement of the layers that meanderingly follow the final instruction vector, plotted as a Feynman matrix vector.

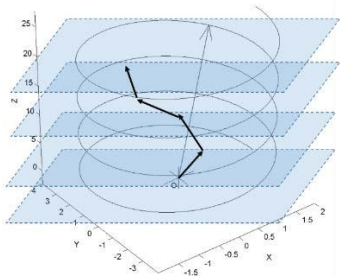


Figure 10: Plotting as a Feynman Matrix Vector

We can see that the vector direction changes with each layer. This can be applied to the generation and change of emotions of an initial-stage artificial ego (virtual ego).

2.6 The Function of Emergence (Induction)

Here, we consider how a machine perceives a person with which it is interacting. For this purpose, we newly defined a new mathematical model called the relationship of "anti". The following table shows this definition.

positive	negative	anti	converse	inverse
+A	−A	A : anti-A	A : converse-A	A : inverse-A
plus minus 0(origin)		Inversion of attributes in the same structure A anti-A converse ⁎ inverse	reversing properties (p→q) - (q→p) A p→q converse-A q→p	conditional negation (p→q) - ($\bar{p}$→$\bar{q}$) A p→q inverse-A $\bar{p}$→$\bar{q}$

Table 1

If in accordance with this definition, we let the state of "anti" be the "inverse of converse" (inversion of attributes), the elements of emotions will be as illustrated in the diagram below.

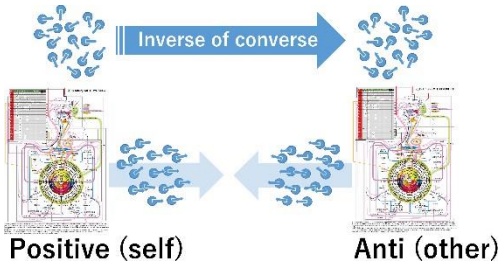


Figure 11: Communication between Positive and Anti

With this, the creation of a robot's drive through inducement with respect to the emotional vector of the counterparty becomes simple mathematically.

A moral and creative autonomous control system for avatars and robots—that is, an artificial ego—can be realized with feedback from anti-fields using the mathematical model we propose. Specifically, this is a model where, regarding a function defined on type U1 that includes as elements all entities that an avatar or robot is aware of under certain circumstances and its opposite, an anti-function defined on type U2 with respect to environment and objects (people and things), when the energy of the function and anti-function are in contention above zero, 1 emerges above U12, which is the opposite of U1 and U2.

If the mathematical model is applied in the following manner, the consciousness and free will of an avatar or robot or what an avatar or robot feels can be defined. More specifically, in a certain framework of thought (circumstances), the domain and range of a function are set so as to include as elements all parameter movements and relationships relating to events that avatars and robots perceive (feel) as being internal, and the domain and range of the anti-function are set so as to include as elements all parameter movements and relationships relating to events that avatars and robots observe as something external. The space combining the domains and ranges of this function and anti-function will constitute the consciousness, or consciousness space, of an avatar or robot. Further, feedback control is applied to the function and anti-function by vectors (free will) implementing control so that the energy of the function and anti-function are in contention at zero above the consciousness space, and after contention, the consciousness space is automatically expanded so that the entities and events up to that point can be captured from the inverse and converse of the consciousness space.

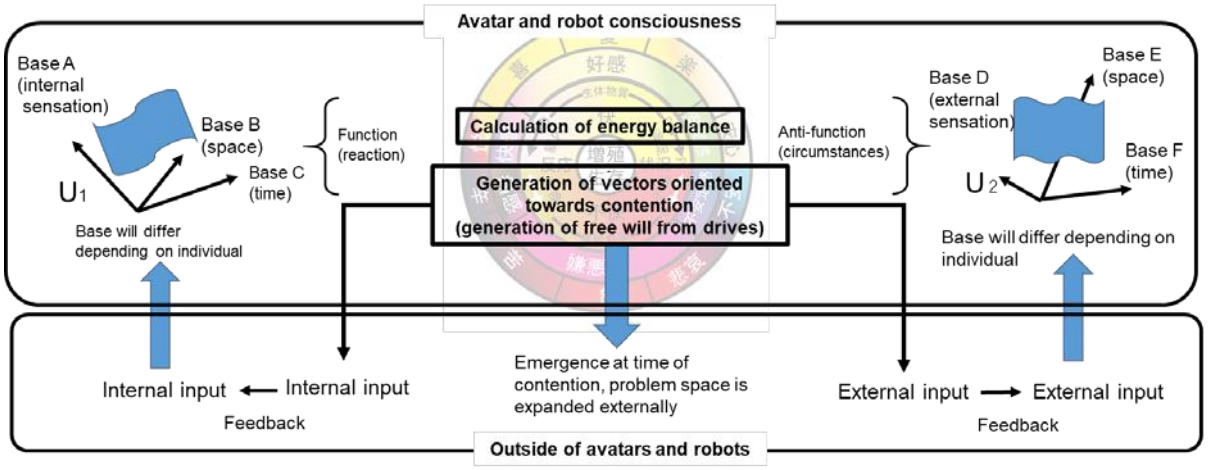


Figure 12: Proposed Model

With this, the concept of "problem" can be defined as an event that impedes the contention between function and anti-function, and the concept of problem solution can be defined as achieving contention between a function and anti-function. Therefore, under our mathematical model, avatars and robots can be realized that have their own free will of trying to solve unfamiliar problems, and when they are solving problems, behavior can be achieved where at the time of problem solving avatars and robot go beyond their framework of thought and attempt to solve the next problem from a larger consciousness space. Further, if avatars and robots can perceive morals as a contention model between the primary free will (drive) that they have internally and cultural knowledge (superego) that they have externally, then it will be possible to write morals control as feedback control of functions by anti-functions; accordingly, if it becomes possible to derive anti-functions through physical experiments, it will also be able possible to implement a mathematical model for moral control. Further, function-based communications can be realized as a mechanism for avatars and robots to transmit the divergence between the function energy and anti-function energy when the two are not in contention. This mechanism seems to show behavior like harmonic oscillators transmitting divergence from the center as waves, but it engages in behavior that differs from wave propagation by harmonic oscillators in two regards, first in that for all individual pieces, there is emergence at time of contention (mapping to U12) and second in that feedback control operates in the consciousness expanded through emergence.

2.7 The Function of Emergence (Phase Transition)

Next, we will explain the mechanism of emergence through phase transition using the formula of the anti-Einstein field hypothesis in the paper discussed above.

$\lim_{nn\to\infty} \text{오} \lim_{nn\to 0} = 0 \Leftrightarrow 1$ shows that with infinitesimal and infinite 오, for both the infinitesimal and the infinite, the limit for countable numbers for both zero and ∞ is the integer 1; for example, with the infinitesimal, the only existence smaller than 1 will be zero. Therefore, this becomes a world of size (continuous quantity) and an argument of $1 \Leftrightarrow \infty$ or $0 \Leftrightarrow 1$. This expresses, as the smallest unit where time and space can be measured, that when mass / 0, then $\lim_{nn\to\infty} \text{오} \lim_{nn\to 0} = 0 \Leftrightarrow 1$. The broken-line red circle in the following diagram corresponds to this. In the paper, within the Schwarzschild radius at this time, the principle of emergence is shown mechanically.

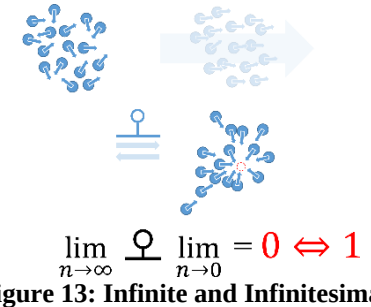


Figure 13: Infinite and Infinitesimal

Through phase transition using this, the class layers in the Feynman matrix discussed above can move freely as in the diagram below. At this time, it is not trial and error but class jumping through phase transition that occurs, and this is used in the principle of emergence.

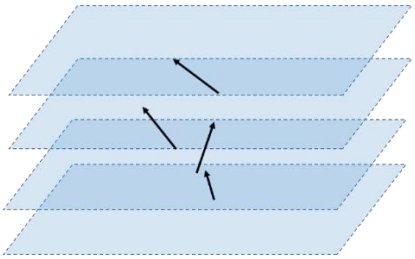


Figure 14: Phase Transition

3. THE "SUPERPOSITION OPERATION" NECESSARY FOR TRUE EMERGENCE

With the infinitesimally small world expressed as the broken-line red circle above, we can think of its structure mechanically as in the diagram below. We then notice that lim and anti-lim are superposed. Instead of multiplication, let us carry out the new "superposition operation".

$\lim_{1\to0}$ $\lim_{1\to0}$

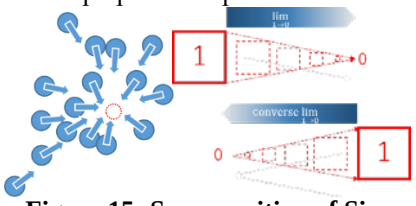


Figure 15: Superposition of Size

"Superposition operation", a technique shown in the following diagram, is a mechanism explained in the above paper, in the section of entitled "Division by zero in an anti-Riemann field".

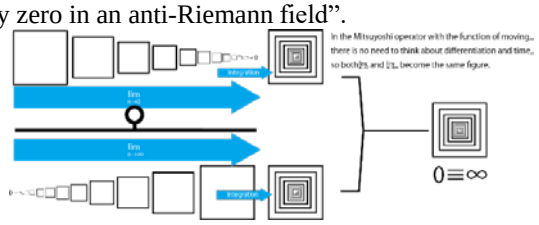


Figure 16: $0 \equiv \infty$

This condition represents phase transition through the colliding energy of ∞ and anti-∞ at the boundary between a null set and anti-null set. We will explain this portion in greater detail as superposition operation.

Let us designate the infinitesimal limit (broken-line red circle) as KU. 1, which is the smallest KU, as expressed by the blue quadrangles. The states from this 1 (blue) to complete nothing (MU) are expressed in the diagram below.

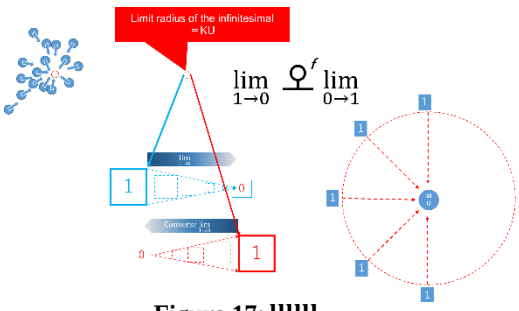


Figure 17: $\lim_{11\to00}$

Next, 1, which comes first from KU, is expressed by the red quadrangles. The states from this MU until 1 (red) are expressed in the diagram below.

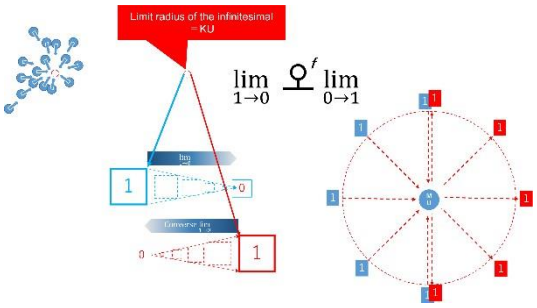


Figure 18: Anti-$\lim_{11\to00}$

At this time, we can see the relationship where, at the left-hand side of the circle, the left half heads from 1 to MU, and the right-hand side goes from MU to emergence at 1. We call this calculation of superposition so that they exist simultaneously the "superposition operation". With this, 1 first emerges at the anti-$\lim_{1\to0}$ field (red semicircular field at left). At the same time, 1 is the $\lim_{1\to0}$ field (blue semicircular field at right) phase transitioning to the anti-field. In the infinitesimally small world, it is ensured that these will occur simultaneously; this is something that we calculated in the foregoing paper, using a Riemann sphere and anti-Riemann sphere.

4. MATHEMATICAL FORMULA OF GOOD AND EVIL

People tend to think of the concept of good and evil through the context of conflict dualism. We thought about the relationship between good and evil.
The following figure shows "good and evil" using division.

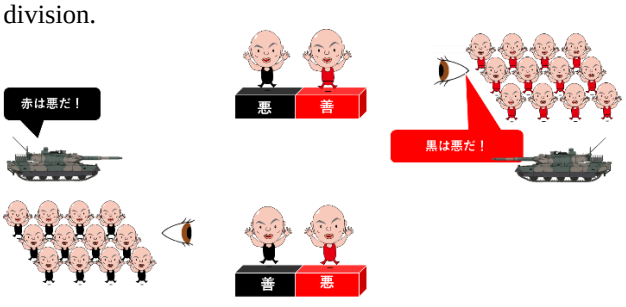


Figure 19: Good and evil
As the figure shows, if what is different from oneself is recognized as an enemy, a conflict

occurs. If we view the recognizable world in terms of “good” or “evil” and consider that good is right and eliminating evil is just and moral, the situation can be shown mathmatically in the following figure.

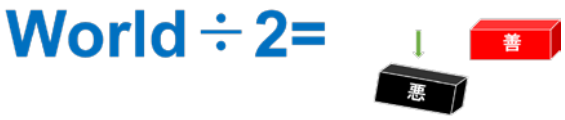


Figure 20: Good and evil expressed with division

However, in this model, if robots and AI consider things by division, the singularity occurs and they will attack people as ordered.
This problem can be compared to a magnet. When a magnet is divided in half, the magnetic polarity of S and N is generated at the respective ends of each half, as shown in the following figure.

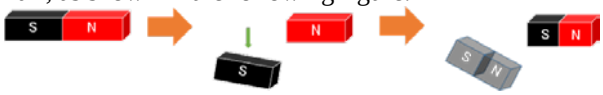


Figure 21: State of good and evil thoughts expressed as magnets

When this is replaced with the previous figure, it appears as below.

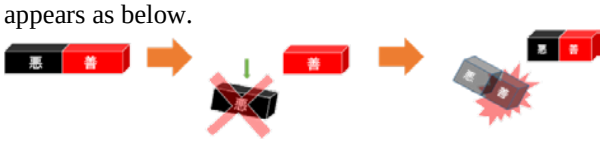


Figure 22: Possiblity of the existence of good in something that is categorised as evil

In this model, the one categorised as evil and then eliminated (killed) has good properties.
Furthermore, considering the physical property of a magnet, the situation will appear as in the following figure.

Figure 23: What if the relationship between good and evil is regarded as a gradation, which is a characteristic of a magnet?

We thought that if robots were capable of recognizing good and evil in a manner analogous to a magnet, the world would change.

The problem becomes an issue of how to realize this problem becomes an issue of how to realize as engineering the “superposition operation” which we posit can be calculated mathematically. Thus we created, similar to “logic gates” in computers, “function-based quantum gates” necessary for artificial ego emergence calculation.

4.1 Dividing by Zero with the Mitsuyoshi Operation (Cut Operation)

There was a need for a symbol that could enable a “concept” not indicating quantity, such as “right” and “left”, to be included in a formula. For example, let us consider an apple being divided into two. With mass of the apple as the reference 1, if this is divided into two equal parts, there will be two pieces with a mass of 0.5. In order to express this state, there was a need to replace the traditional division operator of ÷ with a different symbol for cutting (cut operation operator), namely, the symbol ⨪.

This is Newton’s law of universal gravitation. In 1973, the author tried changing this into a slightly different form.

Figure 24: Newton’s Law of Universal Gravitation

Dividing one apple into two ... One apple became two pieces But with the division operation, the results are as in Figure 25.

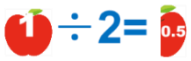


Figure 25: When an apple is expressed with division

÷2= ? Where did the other piece go?

This is a different arithmetic operation from division.
Should we write this?

Figure 26: The Cut Operator, Which is Different from Division

If we do this, as shown in the diagram below, when an apple is divided into two, the parts will not be the same size, there will be a large one and a small one. Let us designate (⨪) as the operation symbol for cutting one apple into two. The means of writing the results (state) is shown below. Figure 4 shows changes in pieces with the cut operation

Figure 27. When an apple is calculated using the cut operation ①

This is also happening. Let's align both ends! And connect both ends with rails! Call the ends

right and left! Let's write in the middle like this!

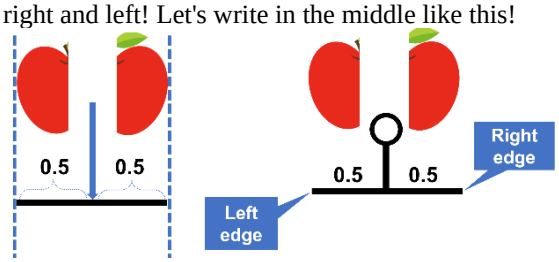


Figure 28. When an apple is calculated using the cut operation ②

This is a sign that the lever moves freely between the left and right ends!

What if the division is slightly off?

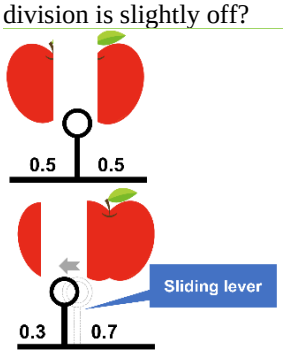


Figure 29. When an apple is calculated using the cut operation ③

We can see that the (오) used here follows the instruction of dynamically moving the right end and left end. If this cut operation operator ÷ and the Mitsuyoshi operator 오 are used to express the state of one apple being divided into two in an equation, this can be written as

1 ÷ 2 = left piece (A) 오 right piece(B)

Equation 1

We notice that the two ends of 오 are in fact not numbers but concepts (consciousness). With this type of calculation, it is possible to easily handle not just numbers and quantities, but also "concepts" such as left and right at the same time within a single equation. Meanwhile, for "quantity", too, with the ratio between the weights of apple pieces A and B (continuous quantities) as x:y, the calculation of the continuous quantity is possible. For example, when actually dividing into two, there are cases where the ratio is not even but is uneven, such as 0.3:0.7.

1 ÷ 2 = 0.5 오 0.5 = 0.3 오 0.7...

Equation 2

In order to be able to write an equation that covers both 0.5 오 0.5 and 0.3 오 0.7, we gave the 오 symbol a slider function. This enabled the boundary between x and y to continuously move. Meanwhile, for "numbers", this can be written:

1 ÷ 2 = A(x) 오 B(y)

Equation 3

Σ ÷ X = A 오 B 오 C〈X-1 = number of 오〉〈Σ = A(x) + B(y) + C(z)〉

Equation 4

This enables a "concept" such as left and right and "numbers" and "quantity" to be written in a single equation, as in Figure 30. The relationship between x and y is like this.

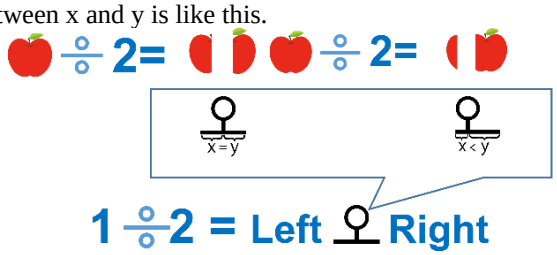


Figure 30: Slider function of Mitsuyoshi operator

÷	÷(cut)	Number of 오
÷X	÷ X	(X-1)
÷3	÷ 3	(3-1)
÷2	÷ 2	(2-1)
÷1	÷ 1	(1-1)=0
÷0	÷ 0	(0-1)=-1

Table 2. Relationship between number of Mitsuyoshi operators and division and cut operation

Let us simplify 1 ÷ 2 = A(x) 오 B(y). First, we stabilize the sum of the respective sides of the =, namely, 1 ÷ 2 and A(x) 오 B(y), as Σ. Next, we turn our attention to length or other continuous quantity (x, y) that is a different aspect of A and B, which are "concepts". If at this time we perform the arithmetic (operation) of moving 오, x and y can be freely changed. For this dynamic function to be able to confine this x, y within Σ, it is necessary only to write:

Σ ÷ X = A 오 B 오 C…

Where (X-1) is number of 오 and X is number of values A, B, C

Equation 5

With this, the "cut operation", which is different from the equal-portion-based "division" and "fractions", was created.

We expressed the means in which robots recognize the world in Mitsuyoshi operators []. The figure below is the model.

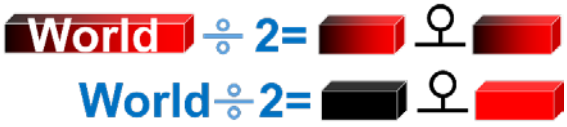


Figure 31: Method of simultaneous calculation of human consciousness with the discrete separation model (digital binary) and continuous model of the natural world

Here, we have confirmed that the gradation model (continuity of the natural world) and separation model (separation of consciousness) are expressed in the same formula.

5. FUNCTION-BASED COMMUNICATION

In order for one (human) operator to control 1,000 or more robots, function quantum bit communications and function-based quantum gates will be effective.
(i) A person's consciousness can be compressed into a function.
(ii) A person's consciousness can be stored in memory as a function, which can be reproduced at any time as a feeling,(iii) These can be achieved for practical use instantaneously with light function-based communication.

Accordingly, function-based quantum gates (patented) will be the protocol.

Quantum jump	Particles	Waves	Gate
KU	MGN(＋)	MGN(−)	MU
$\lim_{m\to 0}$ ○ $\lim_{m\to\infty}$	mass/∞	mass/0	(0≡∞)
Superposition operation	Division by infinity	Division by zero	BH equation

Table 3. Function-based quantum gates will be the protocol.

The following is a comparison between a function-based quantum gate where this computational formula is used as the protocol and an existing logic gate.

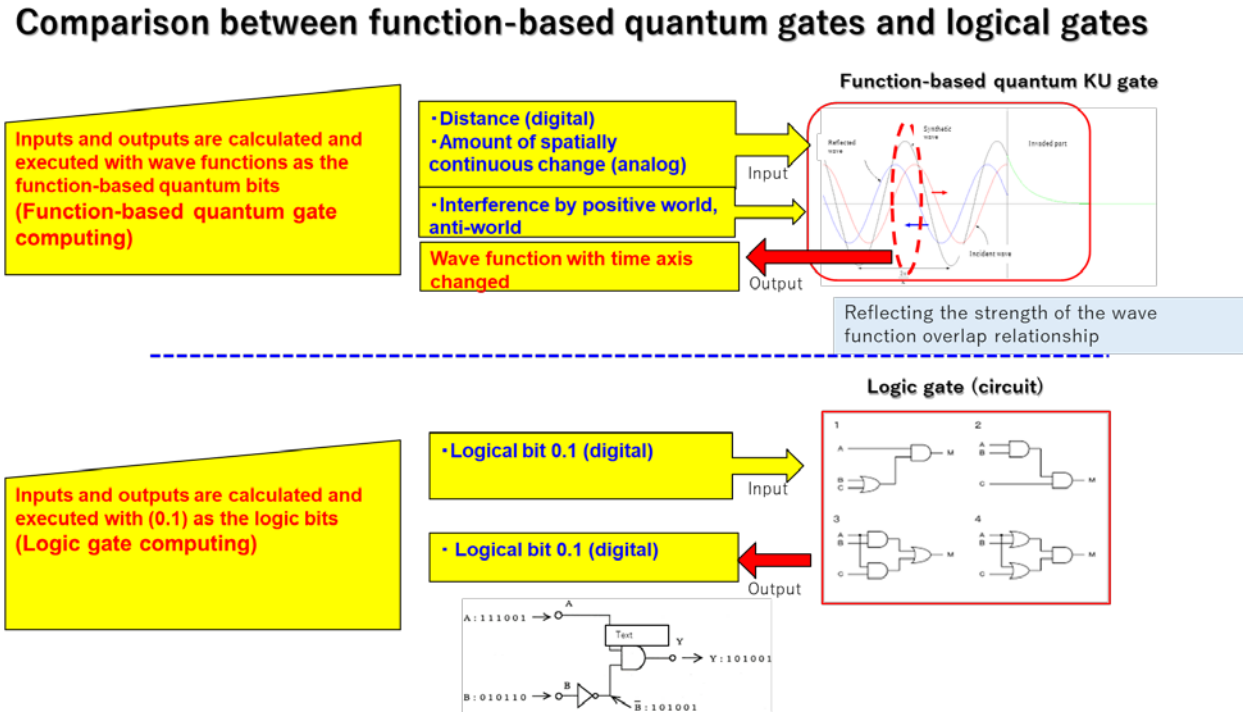


Fig. 37 Function-based quantum gates and Logical gate.

Professor Nakamura of Stanford points out that the Mitsuyoshi operator used here is itself a function-based quantum gate in a quantum state based on an N-dimensional unitary space that is a complex space.

6. MATHEMATICAL COMPARISON BETWEEN AI AND AE.

6.1 Comparison of Mathematical Models

Energy physically corresponds to drive and desire. Accordingly, the states of energy in mathematical models of AI and AE for energy were compared.

6.2 Results

AI, as a natural result of its probability model mathematics, implies entropy, in the statistical mechanics entropy sense of measure of the microscopic "disorder" of a system; therefore, energy is not stable and entropy increases. This happens because, in the transition from a state of equilibrium to another state of equilibrium (phase transition), before and after the transition, system entropy does not decrease but almost always increases. Because of this, AI was fated from the start to be unable to decide a directionality (vector) that represents a clear free will.

However, with the mathematical model used with the AE function-based quantum gate model, the relationship between positive and anti shows the phase transition sum total, so in the positive field entropy decreases, and conversely, in the anti field, sum total balance is maintained by increase or decrease in entropy. This showed that a response (free will) can be given showing the directionality necessary for a judgment as an ego in the positive field.

7. PRELIMININARY EXPERIMENT

The only specific experiment that has been carried out for this research was the public experiment using the initial artificial ego, and the only results that were confirmed were the out-of-control emotions of the robot. How are the emotions of a robot different from the emotions of a person? Is this what is called the consciousness and free will of people? What is this strange behavior of this robot now moving out of control? If this is something that was programmed, then it can't be called the robot's consciousness. These fundamental questions came from a neuroscientist (Dr. Kenichiro Mogi) who attended the experiment. In this paper, we have disclosed these social science concepts of "consciousness", "free will" and "emotion" as mathematical models and algorithms, and have shown their validity using a physics model and a geometric model.

8. DISCUSSIONS

Current neural networks are said to derive answers stochastically, but in reality, they derive concepts of the route to arrive at an answer. However, from the mathematical comparison, we can see that with the AI approach using existing probability models, there is an eternal loop and entropy only increases. The human brain will also repeat loops several times, but after two or three loops will derive a conclusion; if we call this free will and judgment, then the fact that the reduction of entropy and phase transition could be confirmed by the mathematical technique used in the TOE hypothesis using the Mitsuyoshi operator and the function-based quantum gates proposed here, has great meaning. However, change in entropy in the positive field will be change in the inversed anti field. Therefore, what should be done with the anti field? Further, we believe that this will provide a large hint as to whether a system of empathy, where the counterparty is thought about in the context of relationship with the counterparty, can develop as a superego.

9. CHALLENGES GOING FORWARD

Research is needed, using actual prototypes and free conversation avatars, of how the free will and desire generated from the feelings and drives arising emergence of the artificial ego using the function-based quantum gates proposed here are different from those of people. To that end, we are currently building an artificial ego. At this time, as a test corresponding to the Turing test for artificial intelligence, if an artificial ego can provide a problem resolution method for questions that people can't answer, such as "why is it wrong for people to murder?" and "can we have both freedom and equality?" or presents a method for resolving the trolley problem that a human could not come up with or presents other "design or thought that people cannot program beforehand", then we will be able to say that machines possess "consciousness", "free will", "feelings", etc.

10. CONCLUSIONS

10.結論

In this article, we propose an architecture in which robots recognize good and evil. We suggest a new method of recognizing good and evil using Mitsuyoshi operators in a mathematical model, which is expressed in a matrix vector by "free will" and "emotion". We also express "emergence" of "consciousness" in the transition of the matrix vector in the extreme after introducing converse, and the concept of "anti," which is different from converse.

Currently, the only specific experiment of the AE was the public experiment using the initial version of the AE which only has its own emotion. And there were several criticisms to the ambiguity of definitions about "consciousness", "free will" and "emotion". Our proposals above are one of answers to the criticisms.

Further investigation will be necessary to confirm the difference between actual humans and our AE by using robots currently available on the market.

12. REFERENCES

[1] Shunji Mitsuyoshi, "Research on Artificial Ego" Luncheon Seminar at IEEE SMC 2018 (Miyazaki) (2018)

[2] Shigeharu Maeda, "Zusetsu Rinsho Seishin Bunsekigaku", Seishin Shobo (1985)

[3] Rosalind W. Picard, "Affective computing", MIT Press (1997)

[4] Shunjji Mitsuyoshi, "ST ga IT wo Koeru", Nikkei BP (2002), pp. 27-29

[5] The University of Tokyo, "Voice Analysis and Measurement of Pathophysiology", http://www.univ.tokyo/ (refer to July 29, 2020)

[6] Shunji Mitsuyoshi, Softbank, "Profile Feature as Seen" Nature (2016) pp. 198-199.

[7] The Japan Society of Mechanical Engineers, "Kankaku Kanjo To Robot Hito To Kikai No Interaction Eno Chosen", Kogyo Chosakai (2008), pp. 275-307

[8] Stuart Hameroff, Roger Penrose, "Conscious Events as Orchestrated Space-Time Selections" Journal of Consciousness Studies, Volume 3, Number 1, (1996), pp. 36-53.