The Core Principle of Human Intelligence and the Heuristics to Machine Intelligence

Only one author

Abstract

The rapid fluctuation of time and space in the quantum world enables the matter forms in the macrocosm to exhibit complex attributes, which in turn enable our brain to have the abilities of representation, memory, and behavior. In three-dimensional space, the neurons interconnect readily with each other via geometric-topological constructions, then this forms the basis of our consciousness and also supports the principle of material primacy in dialectical materialism. We think that the subjective initiative constitutes the essence of human intelligence, which has shaped our learning and thinking paradigms, and also catalyzes human natural language which naturally supports our subjective initiative and guides our reasoning and planning behaviors. The representation model with the perception-array based on the radial basis function principle, or termed the convolutional principle, is widely used to represent color, voice, shape, smell and taste, even the inspired neural network in deep learning, and accordingly we have done an effective estimate of the maximum number of distinguishable color attributes perceivable by humans. It is known that the stimulation response of the simple neuron in the primary cortex is sensitive to the line-orientation, we further analyze the brain's representation mechanism for a geometric shape which is also the representation models of natural language and mathematical language. We have known that there exist two different brain regions, that respectively analyze what an object is and where it is, so any basic semantics-cluster such as a word or a sentence is represented through a unique neuronal fiber plexus node through the layer-by-layer mapping of an underlying geometric topology just analogous to the node in the neural network in deep learning technique. According to the core concept of category theory, all the semantics have been defined by an elaborate semantics bundle structure in our brain, and we can use them to simulate and interpret the physical world including our behavior. Just as the mathematician Poincare's comment on category theory, it is an art of naming, and we accordingly derive the name-abstract mechanism that has enhanced greatly our abilities of reasoning and planning. Thus, when we perceive a semantics scene, our attention will activate a complex semantics-ripple not a simple conditional reflex, and a behavioral closed-loop termed an intelligence-flywheel can realize the iteration of a semantics bundle structure. We explain why our brain consumes much less power. In the end, from the human intelligence principle, the machine intelligence based on more advanced computer database technique with the semantics-driven or termed the structure-driven mechanism can overcome the bottleneck of von Neumann computer architecture in simulating intelligence.

1 Introduction

The matter attributes in the quantum world, such as superposition states, elusive probability currents and so on, enable them to have extraordinary representation capacity which enables a vast number of aggregated matter forms at larger scales to exhibit their own complex attributes, and also enable life forms to have the representation, memory, and behavior abilities. For example, during the initial stage of life form, the DNA of single cell can store structural information about itself, adapt and evolve within a specific environment through numerous genetic mutations that explore diverse evolutionary possibilities over relatively short life cycles. Over extended periods of evolution, various animals lacking complex brain-based conscious activity emerged. Later, upright-walking humans began making and using tools, greatly increasing the complexity of interactions among diverse matter forms-including humans themselves. Human subjective initiative enabled the development of natural language, and conversely natural language inherently supports the subjective initiative of humans. The neuron-interconnecting structure with geometric topology is capable of representing any information and semantics. In addition, it is well established that two distinct brain regions respectively analyze what an object is and where it is [2], giving rise to an elaborate semantics bundle structure in our brain - a structure that formally defines semantics therein according to the core concepts of category theory. From the foundational principles of human intelligence, we can derive valuable heuristics for the development of machine intelligence.

This paper elaborates several foundational concepts introduced in reference [1]. The paper's subtitles: 1 Introduction, 2 The characteristics of neurons, 3 The color and auditory attributes in geometric space, 4 The shape structure attribute with geometric topology, 5 The trace of visual evolution and visualization, 6 The universal principle of category theory, 7 The behavior with subjective initiative has shaped our thinking-paradigm, 8 Why the brain consumes much less power, 9 The heuristics to machine intelligence from our brain.

2 The Characteristics of Neurons

It is known that each neuron consists of a cell body that contains a nucleus. Referring to figure. 1, branching out from the cell body are a number of fibers called dendrites and a single long fiber called the axon. The axon stretches out for a long distance, much longer than the scale in the diagram indicates. Typically, an axon is 1cm long (100 times the diameter of the cell body),

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

but can reach up to 1 meter. A neuron makes connections with 10 to 100000 other neurons at junctions called synapses. Signals are propagated from neuron to neuron by a complicated electrochemical reaction. The signals control brain activity in the short term and also enable long-term changes in the connectivity of neurons. These mechanisms are thought to form the basis for learning in the brain.

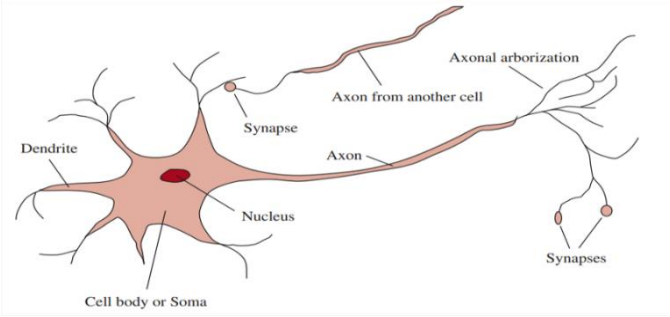


Figure 1. A schematic diagram of a neuron with input and output feature. By defining a nucleus as a node, a specific geometric topology can be represented by the node-interconnecting among neurons.

3 The Color and Auditory Attributes in Geometric Space

3.1 The color representation vector in geometric space

On the retina, three types of cone cells are respectively most sensitive to the light stimulation coming from three different wavelengths 430nm, 530nm, and 560 nm [2]; this is known as trichromatic representation principle. Inspired by the mathematical model used in support vector machine theory, we use the radial basis function to simulate the biological function of these types of cone cells. To better estimate the maximum number of distinguishable colors, we note that the radial basis function with the normal distribution has the largest entropy-requiring the least amount of prior knowledge to generate an optimal simulative shape of the function in a probability space. Thus, the three radial basis functions—also referred to as convolution kernels can be expressed schematically as

$$y_k(f) = a \times exp\left[-\frac{(f-f_k)^2}{2\sigma}\right] (\mathrm{v}),\ k = 1, 2, 3 \qquad (1)$$

where y_k denotes the voltage on the cell membrane, which stems from the stimulation by light of frequency f; a is a constant, and σ is the standard deviation. Here, there are three center frequencies - f_1, f_2, f_3 with $f_1 < f_2 < f_3$.

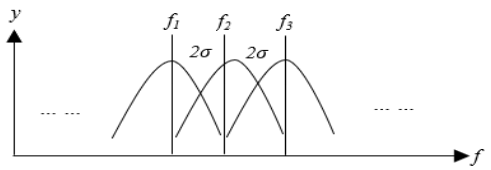


Figure 2. A schematic diagram of the radial basis function representation principle using the normal distribution function for the color attribute.

The support vector machine is a class of optimal linear classifiers in machine learning techniques, and see Figure 2, where the radial basis functions follow the standard arrangement model and can therefore receive relatively stable stimulation energy from each frequency point, so having

$$f_2 - f_1 = f_3 - f_2 = 2\sigma. \qquad (2)$$

We know that, through a voltage ratio $y_1:y_2:y_3$ or a power ratio of y_1^2: y_2^2: y_3^2 which could also be expressed as the form of a feature vector (y_1, y_2, y_3) or (y_1^2, y_2^2, y_3^2), the visual neural system can represent different color attributes. For light of a single frequency f, based on the essential characteristic of this kind of representation model, when f falls within the intervals $[f_1 - 2\sigma, f_1]$, (f_1, f_2), f_2, (f_2, f_3), and $[f_3, f_3 + 2\sigma]$, each specific color attribute can accordingly be represented by

$$y_1, (y_1, y_2), y_2, (y_2, y_3), y_3, \qquad (3)$$

where (y_1, y_2) and (y_2, y_3) denote the joint representation model.

The representation (3) implies that the three types of cells cannot detect the stimulation power less than or equal to $a^2 \times e^{-4}$ which is the value of the y_k^2 when $f = f_k \pm 2\sigma$; this constraint sequentially eliminates the long-tail effect. Note that $a^2 \times e^{-4} / a^2 \approx 0.02$. This value represents the relative detection sensitivity of the cell to the stimulation power, therefore the maximum number of stimulation power representation levels available is 1/0.02 = 50. So, the estimated maximum number of different color attributes we can perceive is 3+4×49=199, which can better represent the visible spectrum as shown in Fig 3.



Figure 3. The visible spectrum is composed of single-wavelength light under normal brightness conditions.

Changes of light brightness can produce the object shadows, which is important in our visual cognition. Thus, they alter the perceived color attributer, and then increases the dimensionality of the color representation space. It is known that the brightness - or termed the grayscale level number the number of brightness levels that the human eye can discern is between 40 and 60. It is not a coincidence. Because no cell has is sensitive to the entire visible spectrum, so the total light energy $\boldsymbol{y}$, that is received through a photoreceptor array which represents the brightness level, equals $y_1^2(f)+y_2^2(f)+y_3^2(f)$, so the relative detection sensitivity remains approximately 0.02 as established early. Accordingly, we adopt 50 as number of the brightness level which can be used to represent the length of the feature vector (y_1^2, y_2^2, y_3^2), and in the natural light-source, noting that 50×50×50×50 = 6,250,000, so it supports the observation that we can perceive millions of color attributes though in the human visual system, $f_2 - f_1 \neq f_3 - f_2$.

Since the human visual system employs another type of radial basis function to capture the changes of adjacent color attributes, consequently the concept of color boundaries emerges. Thus, many patients with color vision deficiency experience difficulties in shape perception [2].



Figure 4. The after-image phenomenon.

As shown in Fig 4, after staring at the bright green disc for about 5 seconds, and then shifting gaze to the adjacent gray circle, observers perceive a faint purplish-red afterimage - the complementary color of green, thus some researchers argued that trichromatic theory cannot fully account for the after-image phenomenon. However, neuroimaging evidence [2] reveals that during such color illusions, brain activity remains elevated in a lower-tier visual area anterior to the V4 region - even after the saturated color stimulus is replaced by gray - indicating sustained neural activation. Therefore, we propose that when gaze shifts to the gray circle, the newly received stimulation elicits a response dominated by change detection: the visual system is selectively sensitive to chromatic contrast.

3.2 The representation of auditory attribute in time-space

In spoken language, a phoneme is the distinctive sound unit that consists of the four acoustic attributes in time and space structure: pitch, intensity, duration, and timbre modulated by our phonatory organ. We can recognize 100 phonemes or so.

4 The Shape Structure Attribute with Geometric Topology

4.1 The compound eye of insects

In order to obtain more information, most insects have trichromatic vision which has a narrower visible spectrum than ours (Menzel, 1979; Chittka, 1996). As an example, *Anax junius* - a dragonfly species that appeared 300 million years ago has 28,672 small ommatidia in a single compound eye and has the smallest inter-ommatidial angle among all known insects, and its visual system synthesizes a mosaic-like image based on a dot-representation model, simulating the most basic geometric structure perceived by insects. With relatively simple cognitive abilities, this group of insects likely established, even in that distant era, a foundational neural mechanism present in nearly all insect brains: two functionally distinct brain regions—one analyzing object identity (“what”) and the other analyzing spatial location ('where ') [2].

4.2 The corneal crystalline eye of humans

A human eye has approximately 260 million photoreceptors, but it has about only 2 million ganglion cells, that is, the efferent neurons of the retina-cells which project to the lateral geniculate nucleus. Most visual information ultimately reaches at the primary visual cortex. In an experiment [2], Hubel and Wiesel found that the output of a simple neuron in the primary visual cortex, whose receptive array receives input from three laterally arranged neurons in the lateral geniculate neurons, is sensitive to the edge orientation which here is termed the orientation attribute of a shorter line segment within a shape. This discovery earned them the Nobel Prize in Physiology or Medicine in 1981.

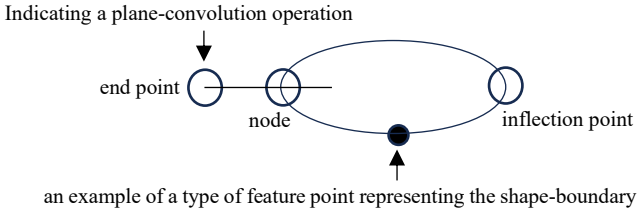


Figure 5. The topology attributes of a geometric shape.

Geometric topology is the essential characteristic of an object's shape, and as shown in Figure 5, the three different types of feature-shape structures – termed respectively the endpoint, inflection point, and node, are the most basic representation-attributes in the geometric structure of an object shape, and for each of which, its structure attribute is defined by the adjacent short line segments related to it, and is different from that of each adjacent short line segment. This aligns with a core principle of category theory: the essence of an object is defined by its relationships to other objects. Thus, at the output layer of the simple neurons in a feature shape structure area. By using the representation method based on a type of radial basis function or also termed the convolution function, the vision system can detect that the output energy strength is different from any adjacent short line segment and sense the existence of a feature point. When those of roughly the same outputs between two of the three feature shape structures along a given short line segment, a linking-mark attribute emerges, and with the attention mechanism our brain assigns the semantics attribute of any area in a visual field. Thus, we refer to the concept of dimensionality reduction for high-dimensional data in manifold theory, and based on the representation attributes with the point, line, and surface the brain's visual system can simulate a geometric shape structure. For instance, during rapid reading, we can still easily perceive the distinct difference between the symbol '**S**' and another symbol derived from rotating symbol '**S**' in three-dimensional space. In addition, referring to Figure. 5, for a circle or an ellipse, our attention

focuses on analyzing the spatial distribution of the four feature points on the far left, far right, topmost, and bottommost.

Our ancestors learned to combine different geometric shapes using syntactic rules in hieroglyphic and symbolic languages to represent and simulate the scenes – including human behaviors, so the semantics presentation consistently exhibits a hierarchical structure based on geometric topology. Here, it is impressive that variations across typefaces fonts show the flexibility of geometric topology and the rigidity of semantics representation based on it.

4.3 The geometric topology and semantics bundle structure

We have known that there exist two different brain regions which respectively analyze what an object interesting is and where it is at [2], and it implies that the basic semantics-cluster such as a word or a sentence is represented through a unique neuronal fiber plexus node arising from the layer-by-layer mapping within a general underlying geometric topology. Just as a letter is represented by a specific geometric topology, a word is represented by combining some letters with the linking-mark in the syntactic sense, as well as a sentence. In brief, the brain does not repeatedly memorize any basic semantics-cluster including a color, a phoneme and so on, and our behavioral activity shapes this elaborate semantics bundle structure. According to the core principle of category theory, that is, the essence of an object is defined by its relationships to other objects, so any semantics-cluster is defined by a semantics substructure related to it, and then we can understand any semantics represented therein and use it appropriately. Therefore, through neuronal message-spikes the semantics information coming from a scene observed by us can diffuse much more effectively, so our brain can better understand and analyze a semantic scene which usually unfolds as an ongoing process, Just as Karl Marx (1818–1883) stated: “Things are universally interconnected and interact with each other; we perceive thing as part of an ongoing process, and this process is about the thing we are perceiving [3]. In addition, we can be aware of a semantics scene being unreasonable, such as when we read a sentence ' An iron ball is floating in the air ', so both correct and incorrect semantics jointly constitute a complete semantics map.

As an example, the word “cat” consists of three letters governed by syntax, and each letter is defined by a distinct geometric topology. The subject here is to represent this word's name with a unique neuronal fiber plexus node, analogous to the node in a neural network in deep learning; its semantics is defined by the related semantics bundle substructure. Here, the three letters syntactically organized are just the abstract of this name, and the abstract of a sentence consists of some words. Generally, any semantics-cluster has a representation subject, i.e., a name expressed by a sentence related to an associated semantics abstract, such as an article. Thus, our brain can use the name-abstract mechanism to concisely annotate a semantics-cluster with a hierarchical geometric structure, then a semantics bundle structure is represented through the interconnections among semantics-nodes, thus this has greatly enhanced our thinking ability. It is impressive that the mathematician Poincare (1854–1912) also once said category theory is 'an art of naming'. In addition, when you see a visual scene, you can say you have seen a rich vision content. However, when you close your eyes, you may only remember some semantic names, concise semantic abstracts, and some salient geometric topologies in it.

Each word can be classified as a noun, a verb, an adverb and so on, thus the concept of semantics-generation enables the brain's analogical reasoning ability, such as the analogy between electric current and water current. Under conditions of insufficient information, fuzzy reasoning is a kind of cognitive paradigm in discontinuous semantics-chains driven by causal relationships, hereby enabling intuitive inference.

5 The Trace of Visual Evolution and Visualization

5.1 The evolution's trace of visual cortex layer

The evolution of the visual cortex layer is based on the mechanism of visual cortical column expansion [2]; that is, the existing visual cortical columns at a given cortical plane originate from the replicative extension of a single ancestral visual cortical column during evolution. For instance, when we view a word 'cat', we can simultaneously perceive information about its color, geometric topologies, constituent letters, and the name of this word form. All these are based on the principle of analyzing geometric topology via parallel processing model, and with the short-time memory mechanism, we can apprehend more structured semantics.

With the attention mechanism, neuronal message-spikes can activate (or 'light up') the associated neuron-nodes representing a specific semantics-cluster, so the structured semantics bundle model inherently possesses the subconsciousness mechanism which also furnishes pathways for our attention shifting. Thus, our attention has activated a complex semantics-ripple not a simple conditional reflex, then we can say that our consciousness has been 'lit up'.

A general stimulus from a scene usually triggers our scene-perception, followed by a semantics-analysis behavior, and possibly by an ensuing behavior - its result may influence future semantics-analysis, so that, a behavioral closed loop or termed the intelligence-flywheel emerges which can realize the iteration of a semantics bundle structure. Karl Marx argued that

all sensory and perceptual functions are interactions between the subject and the object, thus when an experience is closest to the sensation described by empiricists, it is more appropriate to term it 'noticing ' which conveys human agency [3].

5.2 The gestalt principle and visual illusion phenomenon

Our brain sometimes subjectively interprets a visual spatial structure. For example, when viewing a picture in two-dimensional space, our brain sometimes tends to analyze it in three-dimensional space, where the true semantics it represents remains unrevealed. Thus, if we rely solely on experience, we may generate visual illusions and trap ourselves in the dilemma of empiricism. Because the complex macroscopic world we can observe emerges from an even more complex microscopic world, all our brain can do are to attempt to better simulate a semantic scene using geometric topology which is also used in natural language and mathematical language, thus any thinking-result produced by humans or intelligence machines must be verified by practice, it is also an indispensable step in the process of running an intelligence-flywheel.

5.3 Above the geometric topology – semantics and syntax

From the previous analysis, geometric topology, semantics with syntax, and behavior propelled by energy are inseparable in human intelligence. See Fig. 6, before natural language appeared humans could already remember and draw the shapes of animals, and understand the semantics they represent.

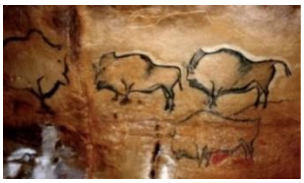


Figure 6. The cave paintings in northern Spain about 40000 years old.

The neuron-interconnecting model based on geometric topology determines that we can only use this type of representation model to simulate the physical world - including our behavior, just as Kant argues in the Critique of Pure Reason. Thus, our brain filters information from the external physical world, and all such information must be represented through geometric topology; this also implies that humans share a common foundational language. In physics, the Maxwell equations in tensor form reveal the intrinsic geometric structure of the electromagnetic field, and its relationship with energy. In addition, the concept of spatial curvature within time-space is also introduced in Einstein's general theory of relativity.

During the past 100,000 years, the number and structure of human brain neurons have not changed. Therefore, if we could revive a human who created the cave paintings in northern Spain about 40,000 years ago, that individual might become a mathematician or physicist.

6 The Universal Principle of Category Theory

6.1 The attribute definition of an object in nature

Here is an example: all the internal attributes within a molecule composed of atoms are in a state of causal closure. Through representation and simulation methods using natural language and mathematical language, its external attributes can be partially revealed by interactions with other molecules. This principle also applies to large molecules and other physical objects defined by us.

6.2 From a computational structure to the relationship between two function forms in algebraic geometry

First, we give an example: for the function $y(x) = x^2$, $y(x_0) = x_0^2$ at $x=x_0$, where the positional attribute of this point can be represented as $[x_0, y(x_0)]$ in a rectangular coordinate system. The first-order derivative at the point x_0 is another attribute relating to this function, which also may be comprehended as the instantaneous rate of change of the function's value. However, intuitively speaking, a rate of change does not exist at a single point in isolation, so we need to use traditional measurement-and-computation method to approximately estimate this change rate. Here, we can construct a computational structure in the neighborhood of x_0, that is

$$y'(x_0) = [y(x_a) - y(x_b)] / (x_a - x_b), \quad (4)$$

where $x_b < x_0 < x_a$, and $x_a \to x_0$, $x_b \to x_0$. Here, what we are actually doing is to use a straight line between x_b and x_a to simulate the geometric structure of this function over the interval $[x_b, x_a]$, and then $y'(x_0)$ can be approximately represented by Formula (4). However, the issues raised by Formula (4) ultimately led to the second mathematical crisis. In order to further compute a higher-order derivative of a function, we need to design a hierarchical computational structure to process more simulated geometric structure information, and as shown in Figure 7, it is

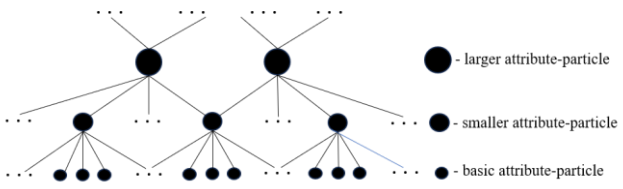


Figure 7. A schematic diagram of a neural network. It is also with a node-interconnecting structure, and each node contains an inner product computation-structure which is the semantic abstract of this node. Any node attribute including that in the interlayer is defined by a specific algebraic geometry structure, each such attribute certainly

should have a specific name, and its confidence level is quantified by a scalar value. The neural network is employed to simulate the geometric attributes of a geometric construction.

the analogous principle to the neural network structure in deep learning, which consists of different attribute-layers or termed feature-layers with radial basis function array or termed the convolution function array, and in which we require to compute the inner product of two vectors in the Hilbert space, in which what we actually do is to compute the similarity between two geometric constructions, one simulating the other (including its attributes) using weights learned during neural network training.

However, the first-order derivative at the point $[x_0, y(x_0)]$ is an objective existence. Actually, there exists another function $y_1(x) = 2x$, which is embedded into the function $y(x) = x^2$ through x_0 and can then represent the first-order derivative value of the function $y(x) = x^2$ at x_0, that is

$$y'(x_0) = y_1(x_0) = 2x_0, \quad (5)$$

where formula (5) realizes the transformation from a computational structure to the relationship between two function forms, and note that formula (5) has lower computational complexity than formula (4). In addition, the famous Hodge Conjecture claims that a complex topological structure can be approximated by combining the geometric constructions raising from polynomial equations.

7 Our behavior with subjective initiative has shaped our thinking-paradigm

From birth, every day everyone constantly needs solve to emerging or future problems, and may propose hypotheses about them, while learning new knowledge. It is as if there is a running intelligence-flywheel in our brain which has been shaping the semantics interconnecting-model which affects our thinking-paradigm for each different thing. Mainly through the reasoning behavior involving in the rational expansion and interpolation of the semantics-chains constructed with different semantics-clusters, we ultimately can complete a semantic jigsaw, where the process by which we engage with any subject with subjective initiative is critical. Through natural language humans can acquire far more knowledge than hominids or other animals, so we can perform complex reasoning and planning about various matters. In addition, at any given moment, our attention mechanism restricts the brain to analyzing only one specific scene, each semantic unit is defined only by a semantics substructure relevant to that scene, so a word may have different semantics in different scenes—for example, 'bank' refers respectively to a riverbank and a financial institution; similarly, 'object' may denote a physical entity or a grammatical constituent. In color cognition, viewing a red rose versus red blood evokes distinct psychological responses.

8 Why the Brain Consumes Much Less Power

Neuron-interconnecting occurs is in three-dimensional space, so our intelligence is based on a semantics interconnecting-structure composed of the semantics-clusters, rather than a computational structure which is separated from computer programs and data, and our brain activates a semantics-ripple by the attention mechanism with the neuron message-spikes. In addition, the threshold voltage required to activate a neuron is very small, around 70 mV; The myelin sheath wrapping around the axon of one neuron forms a current insulator, and the upper frequency of neuron message-spike passing through one neuron is very small such as 1 kHz, thus, there is no electromagnetic interference in our brain, consequently the current flowing through any single neuron is very small, and overall conscious activity consumes only about 20 watts of power.

9 The Heuristics to Machine Intelligence from our Brain

Unlike the human brain, by a vast corpus and image library annotated, large language model (LLM) can analyze the semantic attention-entropy with the statistical prediction and supervised training techniques, and consequently generate a semantic current wrapped in a probability space, then LLM can compress our knowledge obtained and simulate the behavioral paradigms exhibited by humans. However, because LLMs operate entirely in a probability space not in a real physical world, they are prone to increased hallucinations and exhibit diminished reasoning ability.

Noting once again that any semantics-representation must be with the geometric topology form. Along with the development of computer technique, the advanced database technologies, which is based on semantics-driven architectures provides a possibility of simulating our brain which has visualized semantic structure. With the semantic morphism concept of category, a basic semantics-processing unit can operate within the framework of an entire sentence rather than at the level of individual words as in LLM. Therefore, in a specific task, the machine guided by the attention and name-abstract mechanisms can rapidly retrieve those related semantics-clusters with different semantics complexity, then this greatly supports reasoning or planning tasks, while the resulting output can be partially verified against established rational thought chains and domain knowledge. Like the human brain, the machine can readily embed newly formed semantic chains into its architecture, enabling further intelligence evolution.

REFRENCES

[1] Lei Tang, Shi Hui Tang, “The Generation of Semantics in Natural Language and the Formation of Brain Intelligence”, International Journal of Computer Science and Mobile Application, 2024.

[2] Michael S. Gazzaniga, Richard B. Ivry, George R. Mangun, “Cognitive Neuroscience: The Biology of the Mind”, 3rd Edition.

[3] Bertrand Arthur William Russell, “A History of Philosophy”, 1945.

[