Article

Not peer-reviewed version

# From Consciousness to Emotion - On the Intelligent Evolution of Complex Systems

Guangjun Shan *

Posted Date: 1 August 2023

doi: 10.20944/preprints202308.0002.v1

Keywords: strong artificial; intelligence; swarm intelligence; emotion; complex system

*Article*

# From Consciousness to Emotion—On the Intelligent Evolution of Complex Systems

**Guangjun Shan**

Huazhong University of Science and Technology, China; shanguangjun@gmail.com

**Abstract:** Current artificial intelligence (AI) technology is still considered weak AI. However, the progression towards strong AI unavoidably involves grappling with concerns surrounding consciousness, the mind, and creativity. This paper focuses on analyzing consciousness, creativity, and the generation of emotions to explore the possibility of machine-based consciousness, creativity, and emotion. We believe that the emergence of consciousness relies on the complex interplay at different layers within a complex system, rather than wholly relying on the underlying substrates of the system. Thus, it is plausible to consider the existence of machine-based consciousness, creativity, and emotion. Furthermore, swarm intelligence should be recognized as a distinct form of intelligence. Human individuals serve as the foundational substrate enabling the realization of swarm intelligence, with individual intelligence reflecting the projection of swarm intelligence at the individual level. The advanced state of human civilization is primarily attributable to the effect of swarm intelligence. Building upon basic emotions to adapt to environmental conditions, advanced human emotions are a product of the evolutionary process by swarm intelligence. Emotions serve as a pivotal impetus for the formation and development of societies grounded in swarm intelligence. In pursuit of super-AI that surpasses human civilization in the future, nurturing swarm intelligence represents an inevitable trajectory.

**Keywords:** strong artificial; intelligence; swarm intelligence; emotion; complex system

---

## 1. Introduction

With the advent of ChatGPT, a surge of enthusiasm has emerged, instigating a lively discourse on the potential attainment of artificial intelligence(AI) singularity. Functioning as a natural language processing tool propelled by AI technology, ChatGPT engages in dialogues by comprehending and learning human language. Moreover, it interacts with humans within the context of conversations, performing tasks such as drafting emails, generating and interpreting code, and even composing academic papers, all based on human-provided prompts. It is essential, however, to clarify that despite ChatGPT's ability to produce realistic responses and engage in dialogues, it lacks consciousness, emotions, and self-awareness. Built upon statistical and machine learning techniques, ChatGPT learns language patterns and concepts by training on extensive volumes of textual data, enabling the generation of new texts. Nonetheless, it remains incapable of experiencing sensations, emotions, or generating independent thoughts.

The third wave of AI development, which emerged in 2011, encompassed an array of techniques including neural networks and deep learning. This approach heavily relies on human intervention, marking the current phase of AI development as specialized intelligence systems under human control. Conversely, the realization of strong AI necessitates the attainment of autonomous general intelligence with minimal human intervention, a feat that current technology is yet unable to achieve.

When discussing strong AI,it is inevitable to address the definition of "intelligence" itself. What is intelligence? Let's first revisit the debate on the "Turing Test". In 1950, British mathematician Alan Turing proposed in a paper that if a person and a machine were isolated into two rooms, able to communicate but unable to see each other, and if the person could not distinguish whether the conversing entity was a human or a machine, then the machine could be deemed to have achieved human-level intelligence [1]. Critics, however, argue that even if a machine passes the Turing Test, it

 doi:10.20944/preprints202308.0002.v1



doesn't necessarily mean that it possesses intelligence. In response, American philosopher John Searle designed the thought experiment of the "Chinese Room" in 1980. In this experiment, Searle envisages being confined within a room with a book that can handle Chinese characters and grammar. Testers outside the door pass him notes with Chinese sentences through a window. The person in the room can use his book to translate these characters and reply in Chinese. While the testers outside the room may assume the person inside comprehends Chinese, in reality, Searle himself does not understand the Chinese language he is handling [2].

Turing, being a mathematician, inclined toward a functionalist definition of intelligence, while Searle, a philosopher, contemplated intelligence from an intentional standpoint. The crux of their divergence lies in the question of whether intelligence should entail consciousness and a mind. This matter remains contentious to this day and acts as a watershed between current AI (weak AI or specialized AI) and strong AI. Over the past three decades, weak artificial intelligence (AI) has exhibited remarkable advancements, progressively surpassing human performance in diverse domains including chess, expert systems, intelligent search, speech recognition, autonomous driving, facial recognition, and natural language processing. This has repeatedly breached the psychological defense lines that humans have in their understanding of machine intelligence. But the pretentious psychology still compels people to cling to the last bastion of Zion: the belief that machines lack autonomous consciousness and mind. Therefore, despite the immense advancements in machine intelligence, it is unable to alter its fundamental role as a tool serving humanity.

Searle argues that even if a machine is powerful enough to replicate all the functions of the human brain, if it lacks causal reasoning abilities, it can't be said to possess intelligence [2]. Therefore, in the context of strong AI, intelligence requires not only perceptual consciousness, but also self-awareness, the ability to think independently, and to solve problems.

The reason humans consider themselves exceptional is that we believe we possess a soul, which philosophy refers to as the mind. The mind encompasses consciousness, thought, emotion, morality, faith, and more. Humans are distinctly different from their surroundings: we possess consciousness, which not only distinguishes us from our surroundings but also empowers us to make decisions, take action, and shape the world. Within a mere ten thousand years, our extraordinary creativity has fostered a brilliant civilization. We experience a wide range of emotions, including love, hatred, passion, and enmity. Our minds encompass various states such as morality, faith, desire, hope, curiosity, resilience, shame, and more. Can machines possibly possess all of this?

At present, science cannot answer whether AI can possess consciousness and mind. Therefore, exploring the relationship between consciousness, mind, and intelligence from an interdisciplinary perspective can provide a lofty guiding significance for developing strong AI technically.

Before we start the discussion, we need to raise two questions: First, is the belief that machines cannot possess consciousness and mind due to machines' inherent limitations, or is it due to our current understanding and implementation methods of intelligence being insufficient? Second, Must the manifestation of intelligence truly be limited to unfolding solely from our human perspective, and is an anthropocentric definition of intelligence reasonable?

## 2. What is Consciousness

The innate sense of superiority in humans often leads to a resistance to acknowledge that machines, or AI, can possess advanced intelligence. When people say that robots don't have a soul, it's a somewhat vague statement. What aspect of a soul does a robot lack? In terms of causal inference ability, machines have already achieved simple forms of this [3], and they should be able to achieve more complex forms in the future. Human understanding of things has continuously evolved with the deepening of observation and research. Self-consciousness was once widely regarded as one of the most distinctive traits that set humans apart from other animals. However, recent research and experiments suggest that animals such as chimpanzees and dolphins also exhibit a certain

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2023 doi:10.20944/preprints202308.0002.v1



degree of self-consciousness. Before addressing whether machines can possess consciousness and self-consciousness, let's first examine the biological basis of the emergence of consciousness.

The consciousness we usually perceive comes in two forms: perceptual consciousness and self-consciousness. The biological basis of perceptual consciousness is the perception ability of a biological individual from brain neurons to nerve endings, distinguishing itself from the surrounding environment. The reason higher animals, represented by humans, exhibit consciousness is because the body functions of higher animals are highly differentiated and specialized. This high degree of differentiation in body functions necessitates a regulatory mechanism, and the emergence of consciousness fulfills this regulatory function. For example, when a finger accidentally touches a flame, the nerve endings of the finger are stimulated. This stimulus is transmitted to the brain neurons in the form of a bioelectric signal, where the brain converts this stimulus into a sensation of burning pain and instructs the corresponding organs to react. Self-consciousness is more advanced than perceptual consciousness. It has a global understanding of the entire organization and organs of a biological individual, distinguishing not only the individual from the surrounding environment, but also possessing a regulatory mechanism centered on the individual, aimed at seeking advantages and avoiding disadvantages. If AI can receive various input information through sensors, and abstractly process and synthesize this information internally, whether through components or programs, then the self-consciousness of AI may also emerge.

In the study of the emergence of consciousness, some studies attempt to investigate the mechanism of consciousness emergence through a reductionist approach. For example, scientists try to map the circuit connections of brain neurons by slicing biological tissues [4,5], and then study the issues of consciousness and emotion. However, such studies are limited and unlikely to truly clarify the issue of the emergence of consciousness and emotions. A complex system capable of generating consciousness possesses not only complexity, but also hierarchy. Trying to use the physical and biochemical laws from one layer to explain the physical or physiological phenomena in another layer will encounter great difficulties.

This article posits that the emergence of consciousness is a result of the hierarchical nature of complex systems, where information is continually refined and highly abstracted from lower layers to higher layers. For example, the burning sensation caused by a finger touching fire is a highly abstracted result. The individual does not directly perceive how the nerve endings of the finger transmit the signal of burning skin to the brain, or how the brain transforms this signal into abstract information, links this abstract information to the sensation of burning, while also controlling the specific details of swiftly removing the finger from the flame. The brain presents to us not the process of information processing, but the highly abstracted result of information processing (the sensation of burning).

The emergence of consciousness relies on the support of the body's tissues and organs. Consider self-consciousness as an example. Infants, between the ages of 0 to 2 years, do not fully possess self-consciousness, as evidenced by their inability to recognize themselves in the mirror. Around the age of two, as the brain develops and they receive continuous reinforcing stimuli from parents and their surroundings, infants not only begin to recognize themselves in the mirror but also perceive themselves as active agents and can differentiate themselves from others. However, in adults such as patients with Alzheimer's disease, the degeneration of the myelin sheath that protects the nerves prevents signals from being transmitted normally through the nerves. This condition can lead these patients to forget even their own identity, with some losing their self-consciousness. This demonstrates that self-consciousness is not an innate ability of humans but is acquired through physical development and subsequent learning. Furthermore, it can diminish or disappear with the damage to bodily tissues and organs.

While not entirely accurate, we can conceive of the body's organs and tissues as analogous to "hardware", and the process of millions of years of natural selection as a "meta-algorithm". Similar to how computer hardware can be upgraded, biological hardware can evolve through the influence of natural algorithms over millions of years. Similar to a blueprint, modifications to biological hardware

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2023 doi:10.20944/preprints202308.0002.v1



by natural algorithms are passed down from generation to generation through DNA, resulting in the assembly of new "hardware". In the case of individual humans, their "hardware" has evolved over millions of years, forming a complex system that encompasses various biological tissues. The intricate biochemical interactions among different organs and levels serve as the foundation of our consciousness. Figure 1 illustrates the hierarchical structure within this complex system. To give rise to consciousness, the complex system must be layered, operating under simple rules akin to Russian dolls. Various external and internal information flows and is transmitted between different layers within the complex system, with its semantic information highly abstracted. Some of this highly abstracted semantic information manifests in our perceptual consciousness, such as the sensation of burning, or in our self-consciousness programmed to seek benefits and avoid harm.

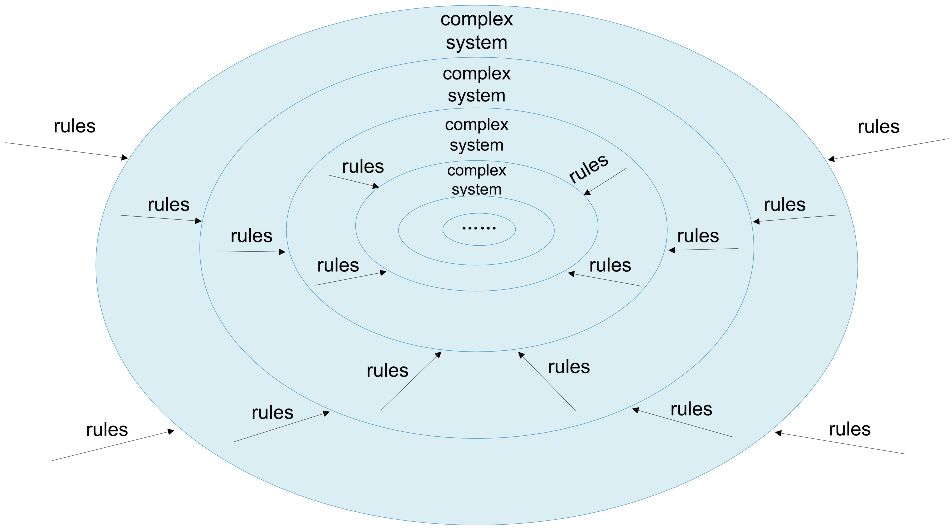


**Figure 1.** Hierarchical Structure of a Complex System Under the Influence of Simple Rules

Building upon the earlier conclusions established in this article, which highlight that the emergence of consciousness stems from the hierarchical nature of complex systems, in which information is continually refined and highly abstracted, we propose a further inference: the genesis of consciousness relies on the intricate interplay within the layered structures of complex systems, regardless of the specific constitutive substratum. Put differently, consciousness can be manifested in carbon-based organisms, silicon-based computers, or any other substrate types, including mixed substrates, as long as the organizational structure of the substrate can sustain the existence of such layered complex systems over the long term. We designate the consciousness arising from silicon-based computer systems as machine consciousness. It is worth pointing out that the complexity of a machine's hardware structure or its computational capabilities alone is insufficient for the emergence of machine consciousness. As previously mentioned, one of the necessary conditions for consciousness to arise is the capacity of the substrate's organizational structure to support the existence of hierarchical complex systems, as depicted in Figure 1. Therefore, if a complex system lacks the ability for hierarchical abstraction of information, it is unlikely to give rise to phenomena such as consciousness or self-consciousness.

## 3. What is Creativity

Another common argument against the potential of machines possessing strong AI revolves around their perceived lack of creativity. This argument often stems from the belief that machines are unable to think, reason, and deduce in the same manner as humans, thus suggesting a deficiency in creativity.

Reasoning ability and associative ability are the two primary characteristics of creativity. However, it is crucial to recognize that all creativity is built upon knowledge. Creative ideas can only emerge based on the foundation of certain knowledge. As an example, the author has personally observed an orangutan in a zoo employing a broken tree branch to access out-of-reach ants. While this action cannot be compared to the vast array of human inventions and creations, its fundamental essence shares common ground.

Referring to Figure 2, let's analyze the process of implementing creativity. The vast majority of creativity originates from a goal or intention (Cause). Some intentions might not be too apparent and are hidden in the subconscious, which we won't analyze here. After establishing an intention, the orangutan must define some abstract rules from the knowledge stored in its brain (prior knowledge) and extract some parts of it. For instance: branches are light, and the hand can pick them up; long objects can extend the length of the hand; ants within the range of the hand can be caught; sturdy railings cannot be moved, etc. Thus, when there is external information input, such as seeing a broken branch, zoo railings, rocks nearby, the orangutan will choose the broken branch as a tool (because the railing cannot be moved, and the stone's length is insufficient) based on prior knowledge and defined rules, extend the "hand's" length to bring the ants within range, and act upon it, hunting the ants (Effect). After this process is complete, whether the result is positive (success) or negative (failure), a new rule will be recorded in the orangutan's brain knowledge database, forming new prior knowledge.

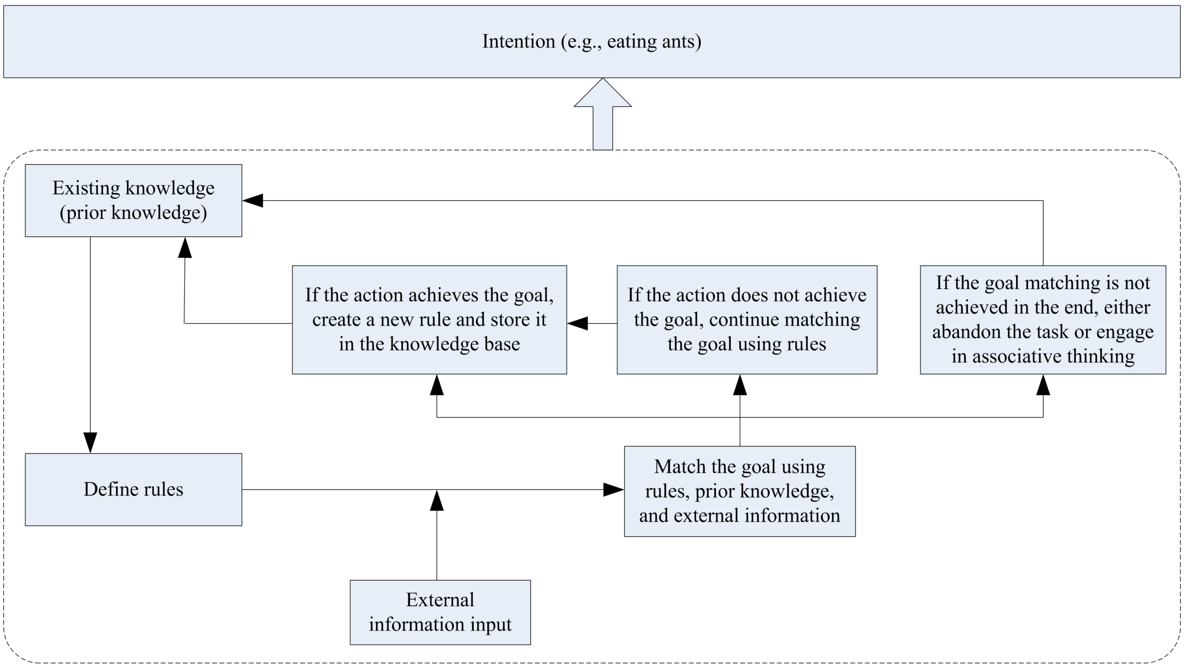


**Figure 2.** Thought and Reasoning Process Flowchart

It is essential to recognize that while thought and reasoning originate from our bodies and brains, the process itself is not consciously perceived by us. If every step of information processing were perceived, it would be unnecessary and exceed the limited capacity of the human brain. Consequently, the information processing involved in creative thinking is highly abstract. We can only perceive limited aspects of the thinking process(causal inference), and the final outcomes(solutions). Due to the brain's lack of conscious perception of the detailed information processing, there exists a sense of mystery surrounding thinking activities, leading to the belief that machines cannot think, reason like humans, and thus lack creativity.

As depicted in Figure 3, Boolean algebra comprises three fundamental operations: AND, OR, NOT. These fundamental operations can be combined to form more complex logical operations. All logical operations, including AND, OR, and NOT gates, consist of "0" and "1" and can be implemented using electronic components such as resistors and capacitors. The highly abstract thinking and reasoning

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2023 doi:10.20944/preprints202308.0002.v1



processes of the orangutan, such as rule definition, applying rules, and prior knowledge to match goals with external information, can be simulated through logical operations. Modern computers fundamentally operate on binary logic operations of Boolean algebra. Looking at the process diagram of thinking and reasoning shown in Figure 2, if a machine can continually incorporate new knowledge or new rules into its existing knowledge base through logical inference with the input of external information, achieving self-updating and iteration, it essentially exhibits typical characteristics of creativity. Whether this creativity originates from natural design or human design does not impact the essence of its existence.

AND

| A | B | $F = A \cdot B$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

OR

| A | B | $F = A + B$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

NOT

| A | B | $F = \overline{A \cdot B}$ |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

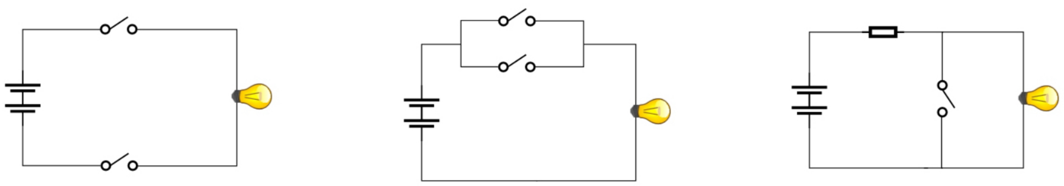

**Figure 3.** Logic AND, OR, NOT operations and their implementation in simple circuits

Apart from reasoning ability, associative ability is another characteristic of creativity. When reasoning fails to explain causality, associative ability can match a plausible explanation based on prior knowledge. This explanation can neither be confirmed nor falsified, but it completes the cycle from cause to effect, which is positively significant for the generation of creativity. Early human religious practices (sorcery) and totem worship originated from this. Due to space limitations, this article will not elaborate further.

Through the above analysis of the process of creativity, we can conclude that mathematical logic can be used to model creativity. In fact, automatic reasoning machines [6] and expert systems are effective attempts to develop AI using mathematical logic. In recent years, compared to the strong momentum of deep learning in the field of AI, automatic reasoning machines and expert systems have relatively stagnated due to their relatively closed applications. However, by continuously enriching and improving the machine's prior knowledge through big data, deep learning, and other technologies and combining it with mathematical logic, it is entirely possible for machines to achieve original creativity. Of course, for machines to reach or even surpass human creativity is a complex system engineering task. Our current capabilities may be insufficient, but with the development of technology, it's only a matter of time before machines possess creativity that we humans either understand or don't understand.

## 4. Swarm intelligence

In the previous section's conclusion, we raised two questions. The analysis of the first question has been completed in the preceding text. Now let us consider the second question: can the manifestations of intelligence truly be limited to our human-defined and imagined perspectives? Is the anthropocentric definition of intelligence reasonable?

Although experiments have shown that animals such as chimpanzees, dolphins, and crows possess a certain degree of self-awareness, self-awareness is still widely regarded as one of the most distinctive traits that differentiate humans from other intelligent entities. However, self-awareness also

 doi:10.20944/preprints202308.0002.v1



imposes certain limitations on our human comprehension of alternative forms of intelligence. 320 AD, the Warring States Period of ancient China, on a bridge over the Haoshui river, a profound debate took place between two eminent philosophers, Zhuangzi and his friend Huizi. The debate centered around the happiness of the fish in the river and how both parties could ascertain the happiness of these fish. This philosophical dialogue can vividly illustrate the cognitive barriers that can exist between distinct intelligent systems.

Zhuangzi observed the fish swimming in the water and remarked, “The fish swims blissfully in the water. How joyful it must be !". Huizi challenged him, saying, “You are not a fish, so how can you know that the fish is joyful ?". Zhuangzi replied, “You are not me, so how can you know that I don’t know the joy of the fish ?". Huizi retorted, “I am not you, and I naturally do not know you. You are not a fish, and it can be completely determined that you do not know the fish’s joy".

The debate reflects the cognitive barriers that exist between heterogeneous intelligent systems due to the lack of effective means of communication. Currently, the definition of intelligence generally relies on human intelligence as the reference point for evaluation. However, a broader understanding of intelligence should encompass not only intelligible systems from a human perspective but also intelligent systems that exist but are temporarily beyond human comprehension, for instance, swarm intelligence.

cIndividual ants exhibit little intelligence and can only respond instinctively to external stimuli. However, ant colonies display a level of intelligence that is vastly disproportionate to that of an individual ant. Ant colonies can navigate obstacles, find food through the shortest path, engage in warfare, construct nests, and nurture their offspring, among other complex behaviors. Previous research has revealed that ants release a chemical substance known as pheromones, which play a crucial role in their foraging process to find the shortest path. When multiple paths exist between the ant colony and a food source, ants tend to make more round trips on the shortest path, leaving behind stronger pheromone trails. After an initial period of random exploration, ants then follow the trail with the strongest pheromone concentration (i.e., the shortest path) to obtain food. Recent studies have further demonstrated the significant role of simple secretion and recognition of pheromones in ant communication and the emergence of swarm intelligence within ant colonies. Researchers have used gene editing techniques to render the ant’s pheromone receptor organs ineffective. As a result, the mutated ants no longer march in a disciplined column (which is why ant colonies can cross obstacles like rivers). The research also found that pheromones regulate ant reproductive behavior. Volatile pheromones secreted by newborn larvae stimulate the production of oxytocin in the adult ants’ brains, which in turn leads to their nurturing behavior towards the larvae [7,8].

An individual ant’s brain consists of approximately 250,000 neurons and cannot be compared to the intelligence of higher animals. However, when ants form colonies, they exhibit highly organized behaviors and accomplish a range of intelligent behaviors typically associated with higher animals, such as nest-building, nurturing offspring, and launching raids for plunder. This article argues that two crucial factors contribute to this phenomenon. The first factor is the physical structure of ants, referred to as their "hardware." Over millions of years, nature has seemingly evolved a “meta-algorithm” for constructing and modifying biological organisms through the laws of nature. The meta-algorithm exhibits the following characteristics: the optimization of biological behavior rules and biochemical parameters by the meta-algorithm is fixed in the current time and space, but it dynamically changes across different time and space. At a specific point in time, organisms constructed based on the “DNA blueprint” possess innate behavioral rules and biochemical parameters. For example, the individual ant’s movement rules and the concentration of released pheromones are already optimized by this meta-algorithm. Thus, ants, as biological entities, provide the substrate for the emergence of intelligence. The second factor is pheromones, which enable communication and coordination among ants within the social structure of the colony. Pheromones transmit a form of symbolic information that is not fully understood by humans, triggering certain innate behavioral rules in ants. This, in turn, gives rise to emergence behaviors in the collective actions of the ant colony, exemplifying a form of intelligence.

Ant colony intelligence is highly inspiring. Even though individual ants lack intelligence, they can serve as the underlying substrate for the emergence of swarm intelligence. In the case of human society, composed of highly intelligent human individuals, what would be the role of swarm intelligence?

For a long time, researchers in AI have focused on studying individual intelligence while overlooking the role of swarm intelligence in human intelligence enhancement. The reason behind this is akin to the debate between Zhuangzi and Huizi, where humans cannot perceive the happiness of fish. Human individuals still cannot perceive swarm consciousness and have not recognized swarm intelligence as a form of intelligence. Section 2 mentions two types of consciousness that we usually experience: perceptual consciousness and self-consciousness. However, there should also exist the third form of consciousness: swarm consciousness. Currently, humans can directly perceive the first two types of consciousness on a physiological level, while the third one is indirectly experienced through belief, social norms, moral constraints, and a sense of national/collective honor.

Why has human civilization reached such an advanced level, at least far surpassing other species on Earth? There are currently two main viewpoints.

Viewpoint 1: The number of neurons in the human brain is greater than in other animals.
Viewpoint 2: The DNA/genes of humans are more advanced than those of other animals.

The number of neurons in the brain is one necessary condition for intelligence, but it is not a sufficient condition. Although research shows that the number of neurons in the human brain is three times that of chimpanzees [9], if we only consider the number of neurons, the elephant's brain has three times as many neurons as humans, why did elephants not become the dominant species on Earth. Even when considering factors such as the connectivity between neurons and the structure of the brain, it seems insufficient to explain why humans, with only a multiple difference in the number of neurons compared to chimpanzees, can generate brilliant civilizations while chimpanzees cannot.

Is there really a distinction between "superior" and "inferior" DNA or genes? The evolution of animals and the genes of living organisms are the results of environmental selection. Those ill-suited to the prevailing environment are eliminated, while the surviving species accumulate genes that have evolved over billions of years. The genome of an octopus is even more complex than that of humans, with a greater number of genes. Therefore, while genes play a role in why humans stand out and become the dominant species on Earth, there are likely other factors at play as well. In essence, DNA and genes can be seen as threshold factors.

The advancement of human intelligence primarily stems from the accumulation of knowledge. It appears that the accumulation of knowledge is much more efficient than the accumulation of genes. Even during the Paleolithic era, which spanned millions of years, and tens of thousands of years ago in the Neolithic era, humans living at that time, who were still hunter-gatherers, were not significantly superior to animals. However, since the birth of human civilization, we have become the absolute dominant species on Earth in just ten thousand years. If we were to place a human child or adult from ten thousand years ago into modern society and provide them with a good education, the possibility of them becoming a social elite would not be significantly different from modern humans. They would also be capable of mastering poetry and quantum mechanics. Let us consider a real-life example. When the British arrived in Australia, the different indigenous populations in Australia were roughly in a transitional period from the Paleolithic era to the Neolithic era, meaning their level of civilization lagged behind modern civilization by tens of thousands of years. Additionally, research indicates that the genes of Australian indigenous peoples became isolated from other human populations on Earth approximately thirty thousand years ago, creating a genetic island [10]. However, David Unaipon, an Australian indigenous man who received an education in a church school during his childhood, demonstrated exceptional abilities in engineering, literature, and several other fields that surpassed the

 doi:10.20944/preprints202308.0002.v1



average person. Currently, David Unaipon's portrait appears on the front of the Australian fifty-dollar banknote. Moreover, such examples are not isolated cases. Can we attribute this person's genes and the development of their brain neurons to a transformative breakthrough that occurred within just a few decades?

Evidently, neither a person's genes nor the number of neurons in the brain cannot undergo significant changes in just a few decades. Thus, we must question what mechanisms are truly at work in facilitating such transformations.

This article posits that the accumulation of knowledge has played a decisive role in the development of human civilization. It acknowledges the positive contribution of brain neurons and DNA to human intelligence but argues that the accumulation of knowledge is a crucial factor in achieving modern civilization. Brain neurons and DNA function more as threshold factors, wherein once individual intelligence surpasses a certain threshold, the accumulation of knowledge becomes the primary determinant. As depicted in Figure 2, the accumulation of knowledge corresponds to the continuous augmentation of prior knowledge. This process relies on a certain group size, as within a group, the rapid exchange of knowledge through communication surpasses an individual's capacity to update personal knowledge. Consequently, from a statistical perspective, the knowledge and skills possessed by an individual within a group represent the average level of intelligence within that group. If the group does not reach a certain size, civilization cannot expand through communication and conflict, hindering the dissemination of intellectual ideas. The "Tasmania Effect" exemplifies this phenomenon. When Tasmania became an isolated island from the Australian continent over ten thousand years ago due to the Earth's changing environment, its small population size and isolated natural environment impeded the transmission and development of civilization, leading to a regression to a primitive level.

Language and writing have played a crucial role in the evolutionary history of human civilization, with writing considered the most significant invention in human history. Through writing, knowledge is no longer reliant solely on oral traditions but can be disseminated and preserved more widely and over extended periods. The cumulative intelligence of generations has accelerated the level of civilization in human society. The question of whether heroes create history or history creates heroes has long been debated. From the perspective of swarm intelligence, it is a process of mutual entanglement and iterative progress. Prominent individuals throughout history are the outcome of swarm intelligence projected onto individuals, and their emergence drives the advancement of swarm intelligence. Thus, even in the absence of figures like Einstein and Newton, it is reasonable to expect the appearance of other brilliant individuals, such as Dinstein and Matton, in history—it is merely a matter of time. The average intelligence level of human individuals is, in fact, the projection of human swarm intelligence onto individuals.

## 5. Emotions and Swarm intelligence

Through the aforementioned analysis, it becomes evident that swarm intelligence, as a form of intelligence, represents a breakthrough in achieving strong AI and warrants further investigation. From a cognitive perspective, the continuous accumulation of prior knowledge serves as the direct cause of intelligence. Groups play a crucial role in facilitating knowledge generation through activities such as competition, cooperation, and communication, while the sharing and preservation of knowledge within the group are achieved through writing and oral traditions. Consequently, the group's prior knowledge expands at a faster rate than that of individuals. In this process, emotions act as invisible driving forces in the formation and development of human swarm intelligence. In turn, the development of human swarm intelligence promotes the evolution of advanced emotional biological functions in individuals.

 doi:10.20944/preprints202308.0002.v1



Emotions have at least the following functions in the formation of groups: the function of connection, introspection [1], and extrospection [2], as well as the regulation of collective behavior.

Connection is a fundamental function served by emotions. To understand human emotions, we can draw insights from the concept of pheromones in ant colonies, as both facilitate the connection between individuals. However, human emotions go far beyond the function of connection. What are emotions and how are they generated? First, emotions are biochemical reactions that act on the organs and tissues of the human body and are perceived as abstract symbols by the human self-consciousness. Therefore, we can conclude that emotions are built upon self-consciousness. Without self-consciousness, emotions would have difficulty fulfilling their functions of introspection and extrospection.

Scientific research has found that certain amine neurotransmitters secreted by the human brain, such as dopamine [11], norepinephrine [12], and serotonin [13], stimulate neurons and are directly related to various emotions unique to humans. The different balance states of these neurotransmitters lead to different changes in human emotions [14]. However, the generation of emotions is not solely determined by neurotransmitters. Other studies have shown that emotions are also influenced by the interconnected structural networks in the brain. According to the triune brain theory proposed by American neurologist Paul D. MacLean, the human brain is composed of the reptilian brain (with a history of 200 to 300 million years), the mammalian brain (with a history of 50 million years), and the neocortex (with a history of 3 million years) [15]. The triune brains represent different stages of human evolution and although they have different functions, they are closely interconnected and mutually influential. The reptilian brain is primarily responsible for maintaining individual survival, including essential physiological functions such as heartbeat, respiration, digestion, excretion, sleep, wakefulness, and bodily sensations. The mammalian brain refers to the limbic system of the brain, primarily including the hypothalamus, thalamus, hippocampus, and amygdala. Its main function is to trigger emotions and process emotional memories. The neocortex is primarily responsible for advanced cognitive abilities such as thinking and planning, helping individuals find solutions to problems and regulate excessive emotions.

The triune brain theory supports the assertion that primitive emotions gradually evolved into advanced emotions as organisms adapted to their survival environment.

Current research attributes the generation of different emotions to different parts of the brain. It is believed that basic emotions (primitive emotions) such as fear and joy are generated in the reptilian brain, while other higher-level emotions are generated in the mammalian brain, and so on. This article holds that in the process of biological evolution, basic emotions arise from an organism's stress responses to the surrounding environment. For instance, when perceiving danger or discovering food, the body releases adrenaline, which elevates heart rate and respiration, increases blood flow, and raises blood sugar levels, thereby enhancing strength and reaction speed. Basic emotions primarily serve to regulate individual physiological and psychological activities. However, whether reptiles and other reptilian organisms possess emotions such as fear and joy as experienced by humans remains an unproven proposition, akin to the debate between Zhuangzi and Huizi. Additionally, this article argues that advanced emotions are differentiated and evolved from the basic emotions, based on adaptive needs of collective living environments. Advanced emotions include likes, loves, hates, jealousies, dislikes, appreciations, loyalties, and more, which reflect individual interactions, group dynamics, and even societal attributes. For example, the mating characteristics of organisms are a natural law, whether for primitive organisms or higher organisms. However, the mating behavior of most animals is clearly influenced more by the secretion of hormones in the nervous system, and after the mating period, they often go their separate ways, with some insects even displaying cannibalistic tendencies where the

[1] Introspection refers to an individual's cognition of his situation in the surrounding environment and his position in the group, and based on this understanding, he adjusts his behavior.

[2] Introspection refers to an individual's cognition of his or her environment and other individuals in the group, and based on this cognition, he or she takes measures to change the environment or affect other individuals.

 doi:10.20944/preprints202308.0002.v1



female consumes the male post-copulation, such as in the case of praying mantises and spiders. This shows that the vast majority of organisms do not possess emotions like love. Conversely, humans have developed love as an emotion derived from their mating instincts, which serves as a biological adaptation to social evolution. The emergence of love as an emotion is driven by the instinctual desire for mating and the expectation of establishing a long-term mutually beneficial community, which helps enhance the survival prospects of both parties. Particularly in a society characterized by a large population, this emotion contributes to social stability and rapid population growth, making it a more sophisticated reproductive strategy.

With the continuous growth of the population, driven by the needs for survival and competition, human populations and societies have gradually emerged. In this process, human individuals have evolved advanced emotions with social attributes, such as care, jealousy, sympathy, loyalty, obedience, dependence, disdain, hatred, and more. It can be argued that the richness and intricacy of human emotions, surpassing those of other animals, stem from the highly complex nature of human social structures. However, the explanation of these advanced emotions cannot be fully accounted for by the triune brain theory alone. Certain emotions, particularly those related to social interactions, do not originate from a specific region of the brain but arise through the collective interaction of the triune brains. With the emergence of advanced emotions, human societies have gradually developed various elements of civilization, including religion, morality, beliefs, social norms, loyalty to the collective, and sense of honor. These elements not only exert extensive and profound regulatory effects on different forms of human swarm intelligence but also serve as indirect manifestations of collective consciousness. Therefore, advanced emotions possess both direct and indirect regulatory functions on human collective behavior.

Figure 4 illustrates the process in which human populations, influenced by factors such as climate, geographical environment, animal and plant populations, gradually evolve into a hierarchical complex system in different layers, such as various societies. It is worth emphasizing that the evolution of complex systems across different layers is likely governed by similar, relatively simple, dynamic rules. Lastly, the emergence of civilization, including aspects such as religion, morality, faith, and sense of honor, can be attributed to the role of emotional factors within swarm intelligence.

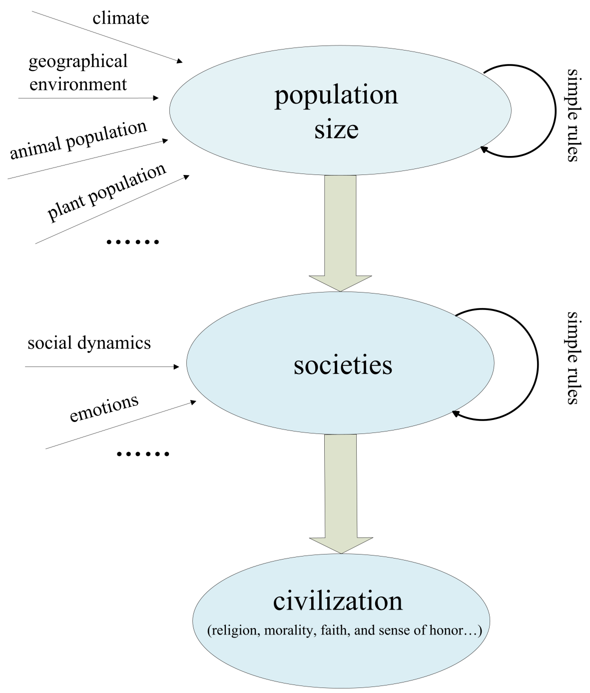


**Figure 4.** The Hierarchical Evolution Process of Human Civilization in Complex Systems

 doi:10.20944/preprints202308.0002.v1



The relationship between emotion and swarm intelligence can be understood from the following two aspects.

Firstly, emotions serve as a catalyst for the formation and development of swarm intelligence in a social context. Both positive and negative emotions play a constructive role in the progress of human society. For instance, showing care for the weaker individuals within one's community not only increases their chances of survival but also fosters empathy, leading to reciprocal support from other members when one faces difficulties. Jealousy, despite being a negative emotion, can stimulate self-reflection and motivate individuals to strive for success in competition with others.

Secondly, emotions are a necessary condition for the advancement of swarm intelligence to higher stages, culminating in the formation of collective consciousness. Swarm intelligence, as a form of intelligence, has the potential to develop into collective consciousness once it reaches a certain stage. Section 4 highlights that due to the physical limitations of individuals, direct experience of collective consciousness is currently unattainable. However, it can be indirectly perceived through various means. For example, religious beliefs, social morals, national identity, and a sense of national pride all objectively reflect collective consciousness. The essential characteristics of living organisms involve continuous self-replication and self-renewal. When a group of individuals forms a social collective bonded by emotions, this collective behaves like a living organism and undergoes constant self-replication. Initially, the replication process follows simple rules, but as these rules are repeatedly applied across different scales of population and exhibit distinct sociodynamic characteristics at various levels, the collective can generate complex social structures through relatively simple patterns. At this stage, new features emerge that correspond to a complex society, namely the aforementioned collective consciousness. Collective consciousness can regulate the behavior of individuals within the group through introspective norms such as belief systems, moral values, and sense of honor, enabling the collective to better utilize its surrounding environment and gain competitive advantages.

## 6. Emotion and Strong Artificial Intelligence

The preceding analysis highlights that the formation of emotions is a necessary prerequisite for the development of swarm intelligence. Currently, the field of machine emotion modeling is in its early stages. The common approach in emotion computation and recognition involves collecting physiological signals or facial expressions from the human body using external devices such as wearables and cameras. Subsequently, physiological and emotional features are extracted. Emotion pattern recognition algorithms are then employed to distinguish and recognize emotions, and corresponding emotional expressions are manifested through facial expressions, body gestures, language, voice, and other modalities [16,17]. Thus, the current machine emotion modeling involves simulating and extending human emotional behavior or expressions, resulting in embodied emotional machines that are incapable of spontaneously generating primary emotions.

An AI system devoid of emotions does not attain the status of strong AI. The realm of strong AI inherently encompasses dimensions such as consciousness, creativity, and emotions. As elucidated in the earlier discourse, both consciousness and creativity, as well as emotions, can be achieved not solely by biological organisms but also through the amalgamation of machine hardware and algorithms. As mentioned in Section 5, the two fundamental attributes of living organisms are self-replication and self-renewal. To some extent, a machine's ability to self-renew can be seen as an expression of creativity. Figure 2 illustrates that machine creativity lies in its capacity to deduce novel patterns and knowledge from external information, subsequently integrating them into its knowledge repository to form new prior knowledge. Categorized by the extent to which machines can realize self-replication and self-renewal, the classification of strong AI can be divided into four levels, as depicted in Figure 5. It is evident that the higher the assigned level, the greater the potential for intelligent progression.

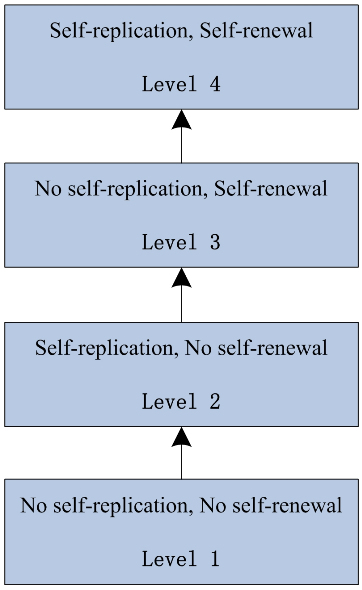


**Figure 5.** Strong AI Hierarchies

This article argues that achieving general strong AI, specifically in humanoid robots, requires addressing the issue of machine consciousness before realizing machine emotions. Without the attainment of machine consciousness, machine emotions cannot possess the subjective experiences and sensations of humans. They cannot generate self-driven behavior from introspection or extrospection and fail to achieve genuine emotional communication between machines and humans. Currently, effective modeling of consciousness remains elusive for humans. As discussed in section 2, consciousness arises as a highly abstract symbol obtained through the hierarchical extraction and abstraction of information within complex systems, such as signals from specific neural fibers. Various external and internal information flows and exchanges between different levels of complexity within the system, with semantic information being highly abstracted. Some of these highly abstracted semantic information manifests as our perceptual consciousness (e.g., the sensation of pain) or self-consciousness (e.g., the concepts of "you", "me", and "others"). Perceptual consciousness is not fixed even for living organisms. For instance, cockroaches naturally prefer to consume food with a sweet taste. However, after years of exposure to sweet-tasting poisoned bait, generation after generation, the taste receptor neurons of cockroaches reclassify the previously sweet taste as bitter, leading them to refuse any food containing sugar [18]. This demonstrates that consciousness itself is, to some extent, merely an abstraction whose interpretation lies within the cognitive system. Self-consciousness can also be subject to errors. In the experiment of the rubber hand illusion, participants often perceive the rubber hand as their own and may exhibit erroneous perceptions [19]. This implies that our strongly held self-consciousness may not be entirely genuine. Understanding the highly abstract symbol of consciousness likely requires a cognitive system for interpretation. Thus, machine consciousness needs to be combined with a cognitive system to generate consciousness in the context of human beings. There is no substantial difference as to whether the cognitive system is the product of nature or of human beings.

Furthermore, existing emotion computation and recognition techniques are insufficient to support the generation of innate emotions in machines. The mechanisms behind emotion generation are exceedingly complex. For instance, even for the same event, individuals may experience different types and degrees of emotions. Some may display great joy, while others may exhibit only moderate happiness or even sadness. Emotions possess relativity, subjectivity, and hybridity (the simultaneous presence of two or more emotions within an individual). To simulate the generation of primary emotions, more suitable technologies and mathematical methods need to be developed.

 doi:10.20944/preprints202308.0002.v1



As mentioned, swarm intelligence serves as a breakthrough for achieving strong AI. Emotions, particularly advanced emotions, are evolutionary products of human society. From the perspective of promoting civilization evolution, the role of swarm intelligence surpasses individual intelligence. Individual intelligence represents the projection of swarm intelligence onto individuals. Following this line of reasoning, if we aim to develop strong AI, we must delve deeper into the development of machine emotions and establish emotional interactions between machines and either machines or humans.

Currently, the thriving field of AI is predominantly driven by deep learning techniques, represented by convolutional neural networks and recurrent neural networks. These techniques mimic the structure of human neurons and train the parameters between neurons using error backpropagation. Through the use of multi-layered neural networks, these techniques extract highly abstracted data features or semantic information. However, swarm intelligence follows a different technological trajectory. Firstly, the formation of swarm intelligence is based on a few very simple rules, with individuals interacting continuously and forming complex social structures under the influence of external factors, such as the environment. Individuals begin to specialize and stratify. Therefore, swarm intelligence is better suited to addressing tasks that are multi-objective and hierarchical, making it more applicable to the generality of strong AI. Secondly, swarm intelligence is decentralized, with each individual making a corresponding contribution to the task. The failure of a few individuals does not result in the overall failure of the task, making swarm intelligence more adaptable to distributed and networked environments. Thirdly, by incorporating emotions, swarm intelligence can achieve faster accumulation of prior knowledge, longer-lasting collective memory, and the emergence of powerful swarm intelligence through interactive inspiration compared to individual intelligence. Currently, the application of swarm intelligence is still in its early stages, with algorithms such as ant colony optimization and particle swarm optimization representing shallow implementations. The powerful emergence effects of swarm intelligence have yet to be fully realized. There is currently a lack of in-depth mechanistic research on swarm intelligence, necessitating a meticulous and comprehensive analysis of the mechanisms underlying its emergence in the future. We believe there are several aspects that should be considered in the development of swarm intelligence technology:

1) The current focus of deep learning techniques is on the connections between neurons, while future swarm intelligence technology should emphasize the connections between intelligent agents. Intelligent agents, as the fundamental substrates of swarm intelligence, will possess structures and functional requirements that are more complex than individual neurons.

2) The design and implementation of machine emotion mechanisms are pivotal to the realization of swarm intelligence technology. Further in-depth research is necessary to investigate the psychological, biological, and sociodynamic mechanisms of human emotions, and based on these findings, abstract machine emotion mechanisms can be designed.

3) The generation of machine emotions will enable machines to participate in higher-level human intellectual activities such as scientific theory research, technological development, artistic creation, and more. This will lead to the formation of human-machine swarm intelligence, driving accelerated development in the field of science and technology within human society.

## 7. Conclusion

Similar to individual intelligence, swarm intelligence is a complex system. Without the evolution from individual intelligence to swarm intelligence, crucial elements such as emotions, language, and writing, which are vital to the development of human swarm intelligence, would not exist. Given our current technological level, achieving human-level strong AI may still be beyond our reach. However, studying the mechanisms and principles of swarm intelligence, developing new AI technologies, and integrating them with current specialized or weak AI technologies can still have a significant impact on our technological progress, akin to a nuclear-level effect.

Another thought-provoking question arises: when we achieve strong AI, will there still be differences between AI and human intelligence? If we continue to design AI based on a deterministic approach, where all machine intelligence makes the same decisions and experiences the same emotions when given the same internal and external parameters, it would differ from the fluctuating behaviors and emotions commonly observed in humans. Therefore, if we desire to design AI that is similar to humans and not so "perfect", it would be necessary to incorporate randomness and uncertainty factors into strong AI.

An ensuing ethical issue is whether machines can still be referred to as "it" if they possess emotions similar to humans.

In the next 100 years, there is a high likelihood of witnessing profound transformations in human evolution. It is very likely that our existence will persist and evolve in ways that are presently beyond our imagination, encompassing not only our intellectual capabilities but also our physical forms. Despite previously believing in our free will, when faced with immense temptations such as eternal life, we will eventually realize that individual perseverance, after spanning several generations, cannot alter the ultimate outcome of collective choices. That is to say, it is highly likely that our descendants will evolve in a manner that we are unwilling to accept in the present.

## References

1. Turing, A.M. Computing Tachinery and Intelligence. *Mind* **1950**, *49*, 433–460.
2. Searle, J.R. Minds, brains, and programs. *Behavioral and Brain Sciences* **1980**, *3*, 417–424. doi:10.1017/S0140525X00005756.
3. Cui, P.; Athey, S. Stable learning establishes some common ground between causal inference and machine learning. *Nature Machine Intelligence* **2022**, *4*, 110–115.
4. Lavazza, A. 'Consciousnessoids': clues and insights from human cerebral organoids for the study of consciousness. *Neuroscience of Consciousness* **2021**, *2021*, niab029.
5. Smith, K. THE BRAIN-CIRCUIT EVEN THE SIMPLEST NETWORKS OF NEURONS DEFY UNDERSTANDING. SO HOW DO NEUROSCIENTISTS HOPE TO UNTANGLE BRAINS WITH BILLIONS OF CELLS? *Nature* **2017**, *548*, 150–153.
6. Morgan, J.L.; Lichtman, J.W. An individual interneuron participates in many kinds of inhibition and innervates much of the mouse visual thalamus. *Neuron* **2020**, *106*, 468–481.
7. McKenzie, S.K.; Winston, M.E.; Grewe, F.; Vargas Asensio, G.; Rodríguez-Hernández, N.; Rubin, B.E.; Murillo-Cruz, C.; von Beeren, C.; Moreau, C.S.; Suen, G.; others. The genomic basis of army ant chemosensory adaptations. *Molecular Ecology* **2021**, *30*, 6627–6641.
8. Trible, W.; Olivos-Cisneros, L.; McKenzie, S.K.; Saragosti, J.; Chang, N.C.; Matthews, B.J.; Oxley, P.R.; Kronauer, D.J. Orco mutagenesis causes loss of antennal lobe glomeruli and impaired social behavior in ants. *Cell* **2017**, *170*, 727–735.
9. Benito-Kwiecinski, S.; Giandomenico, S.L.; Sutcliffe, M.; Riis, E.S.; Freire-Pritchett, P.; Kelava, I.; Wunderlich, S.; Martin, U.; Wray, G.A.; McDole, K.; others. An early cell shape transition drives evolutionary expansion of the human forebrain. *Cell* **2021**, *184*, 2084–2102.
10. Malaspinas, A.S.; Westaway, M.C.; Muller, C.; Sousa, V.C.; Lao, O.; Alves, I.; Bergström, A.; Athanasiadis, G.; Cheng, J.Y.; Crawford, J.E.; others. A genomic history of Aboriginal Australia. *Nature* **2016**, *538*, 207–214.
11. Badgaiyan, R.D.; Fischman, A.J.; Alpert, N.M. Dopamine release during human emotional processing. *Neuroimage* **2009**, *47*, 2041–2045.
12. Barandouzi, Z.A.; Lee, J.; del Carmen Rosas, M.; Chen, J.; Henderson, W.A.; Starkweather, A.R.; Cong, X.S. Associations of neurotransmitters and the gut microbiome with emotional distress in mixed type of irritable bowel syndrome. *Scientific Reports* **2022**, *12*, 1648.
13. Pourhamzeh, M.; Moravej, F.G.; Arabi, M.; Shahriari, E.; Mehrabi, S.; Ward, R.; Ahadi, R.; Joghataei, M.T. The roles of serotonin in neuropsychiatric disorders. *Cellular and Molecular Neurobiology* **2022**, *42*, 1671–1692.
14. Wang, F.; Yang, J.; Pan, F.; Ho, R.C.; Huang, J.H. Neurotransmitters and Emotions, Volume II. *Frontiers in Psychology* **2022**, p. 2860.

15. MacLean, P.D. *The triune brain in evolution: Role in paleocerebral functions*; Springer Science & Business Media, 1990.
16. Fu, Z.; Zhang, B.; He, X.; Li, Y.; Wang, H.; Huang, J. Emotion recognition based on multi-modal physiological signals and transfer learning. *Frontiers in Neuroscience* **2022**, *16*, 1000716.
17. Cugnata, F.; Martoni, R.M.; Ferrario, M.; Di Serio, C.; Brombin, C. Modeling physiological responses induced by an emotion recognition task using latent class mixed models. *Plos one* **2018**, *13*, e0207123.
18. Wada-Katsumata, A.; Silverman, J.; Schal, C. Changes in taste neurons support the emergence of an adaptive behavior in cockroaches. *Science* **2013**, *340*, 972–975.
19. Jalal, B.; McNally, R.J.; Elias, J.A.; Potluri, S.; Ramachandran, V.S. "Fake it till You Make it"! Contaminating Rubber Hands ("Multisensory Stimulation Therapy") to Treat Obsessive-Compulsive Disorder. *Frontiers in Human Neuroscience* **2020**, *13*, 414.