UDC: 681.3 https://doi.org/10.15407/jai2024.03.032

V. Yashchenko
Institute Problems of Mathematical Machines and Systems of the NAS of Ukraine, Ukraine
40, Akademika Glushkova avenue, Kiev, 03680
vitaly.yashchenko@gmail.com
http://orcid.org/ 0000-0001-9396-6581

ARTIFICIAL INTELLIGENCE: FREE WILL, SELF-CONSCIOUSNESS AND ETHICS

Abstract. The article discusses key aspects of artificial intelligence creation, including issues of free will, self-awareness and ethics. The focus is on the neurobiological basis of consciousness, in particular the structure and functions of the new cerebral cortex, as well as the mechanisms of recognition, memory and prediction, which are important for modelling cognitive processes in artificial systems. The paper discusses the role of neural networks in reproducing cognitive functions, such as perception and decision making, and presents modern approaches to training neural networks. A separate part of the paper is devoted to the issue of modelling self-awareness and subjective experience in artificial intelligence and how realistic it is to create self-aware machines. Ethical issues of artificial intelligence creation are at the centre of the discussion, including the topics of the rights of self-aware machines, their responsibilities and their role in society. The article considers the possible social consequences of the emergence of artificial personalities, the need to develop new legal frameworks and legal protections for such beings. It also discusses the problem of free will in the context of both biological and artificial systems, citing experiments and philosophical theories that question free will as a phenomenon. It concludes that the creation of artificial intelligence has great potential, but requires careful ethical and legal analysis to ensure the harmonious integration of artificial persons into social and legal structures.

Keywords: Artificial intelligence, free will, self-awareness, neural networks, digital personalities.

Introduction

The purpose of this article is to investigate the main aspects of artificial mind creation affecting free will, self-consciousness and related ethical issues. The article discusses the possibilities of reproducing consciousness and self-awareness in artificial systems and the correlation of these concepts with biological processes in the human brain. Artificial consciousness is one of the key problems of modern science and philosophy, located at the intersection of neurobiology, artificial intelligence and ethics. With the development of neuromorphic computing technologies that attempt to reproduce the structures and processes of the brain, the question arises: is it possible to create a system with self-consciousness and capable of conscious actions similar to the human mind? If so, should such an artificial consciousness be considered a person with its own free will? The article also analyses experimental studies that question the existence of free will, which raises profound philosophical and ethical problems. The emergence of artificial intelligence has the potential to change both our understanding of human nature and legal and social norms, which requires special attention to the moral aspects of creating self-aware artificial persons.

Fundamental bases of brain function

The human brain is an extremely complex biological system that performs a wide range of cognitive and motor functions. Dividing the brain into different regions allows us to understand how the different components of this system work and which neural structures are responsible for the complex processes of perception, thinking, decision-making and self-awareness. One of the most significant parts of the brain associated with higher cognitive functions is the new cortex, or neocortex. This part of the brain represents the pinnacle of evolutionary development and is considered key to human intelligence and self-awareness.

Structure and function of the new cerebral cortex

The new cerebral cortex (neocortex) is a layer of tissue that covers the large hemispheres of the brain. The neocortex is responsible for higher cognitive functions such as abstract thinking, planning, perception, attention, memory, language and decision-making.

It has a unique structure consisting of six layers, each with specific functions. These layers include pyramidal neurons, star cells, and interneurons, which together create complex networks for processing information. The basic principle of the neocortex is to transfer information between layers and process sensory signals from the outside world, as well as integrate information to make informed decisions and form complex thoughts. In each part of the neocortex, specialised areas (zones) perform different functions. For example, the visual cortex is responsible for processing visual information, while the prefrontal cortex is involved in decision-making and planning. The connections between neocortex areas provide a complex integration of sensory data with internal cognitive processes. It is the interaction between the different cortical areas that allows a person to perceive the world around them, adapt to change and respond to external stimuli. The neocortex plays a key role in the processes of consciousness and self-awareness. It not only provides the ability to perceive external signals, but also participates in reflection and the formation of complex thoughts. These features make the neocortex a critical object of study in the development of artificial intelligence and systems that seek to replicate the human mind. The application of such structural and functional principles in AI may bring us closer to creating machines capable of self-awareness.

Mechanisms of recognition, memory and prediction

The mechanisms of neocortex functioning include not only the perception of information, but also its memory, recognition and prediction of future events. These processes are based on synaptic plasticity - the ability of neurons to change the strength of their connections in response to experience and learning. It is synaptic plasticity that underlies learning and memory in biological systems.

Pattern Recognition

One of the most important functions of the neocortex is pattern recognition. When a person sees, hears, or feels, the sensory signals enter the cortex where they are processed to create a conscious perception. The cortex uses both information from current sensory data and stored memories to identify and interpret the signals. For example, the visual cortex recognises faces, objects and movements, while the auditory cortex recognises speech sounds and music. These processes are closely related to the human cognitive capacity for association, which allows linking new images to already known ones. From the perspective of artificial intelligence, pattern recognition is one of the most important areas that is being actively developed thanks to neural networks. Modelling these mechanisms using deep neural networks helps to build machines capable of image processing, speech recognition and other sensory data. However, despite the successes, existing models still fall short of the human ability to generalise and associate.

Remembering

Remembering information is another critical process that is controlled by the neocortex. The storage of information in the brain occurs through long-term and short-term memory. Long-term memory allows a person to retain knowledge for years, while short-term memory (or working memory) helps retain information for a few seconds or minutes. Neurons in the neocortex play a key role in these processes, maintaining information through strengthening or weakening synaptic connections between neurons. The process of remembering usually involves multiple repetitions and the strengthening of neural pathways that become stronger with each new repetition of information. This process is called long-term potentialisation, which is the basis of long-term memory. Developing AI capable of remembering and learning requires the creation of models that reproduce this process. Modern machine learning methods strive to create models that can adapt to new data and retain information, but they cannot yet rival biological systems in flexibility and depth of memory.

Prediction

An important human cognitive function is the ability to predict. The neocortex is actively involved in predicting future events based on previous experiences. This occurs through the analysis and processing of patterns that the brain extracts from the environment and its own experience. For example, the visual cortex predicts the movements of objects based on their speed and direction. The brain uses synaptic plasticity and associative learning to make successful predictions. These mechanisms allow the neocortex to adapt to changes in the environment and make educated guesses about the future based on previous observations.

Artificial intelligence is also actively using prediction concepts through deep learning models and recurrent neural networks. Such systems are trained on large amounts of data and can predict outcomes based on previously input information. However, further developments in simulation of neural processes and cognitive functions are required to create systems with true predictive potential approaching that of humans. The fundamentals of brain function, especially the role of the neocortex, are critical to understanding how human consciousness and cognitive functions work. The neocortex performs key recognition, memory and prediction tasks that underpin human thinking. The study of these processes is helping in the development of artificial intelligence that seeks to replicate the complex functions of the brain. Creating an AI that will be capable of not only learning, but also prediction and awareness requires a deeper understanding of the neural mechanisms that underlie the workings of the human brain.

Neural networks and their role in cognitive processes

Neural networks are one of the key tools of modern artificial intelligence that mimic the workings of the human brain. Originally developed on the basis of biological neurons, artificial neural networks (ANNs) have become the basis for the creation of powerful models capable of solving a wide range of tasks, from pattern recognition to time series prediction. These systems have greatly influenced the development of cognitive sciences and have helped to investigate the workings of the human mind. This section will discuss the types of neural networks, their relationship to cognitive processes, learning mechanisms, and applications in the cognitive sciences.

Types of Neural Networks

Artificial neural networks can be classified into several types depending on their structure and function. Each of these networks is used to solve different tasks and plays an important role in creating artificial intelligence and modelling cognitive processes.

Single-layer neural networks are the simplest models consisting of a single layer of neurons. Input data is fed directly to the output neurons through weighting coefficients. Each neuron has an activation function that converts input data into output values. At the beginning of training, random weights are assigned to each neuron, which are adjusted during training to minimise the error. The main training method is gradient descent, which adjusts the weights based on the error. Such a network can perform simple classification tasks, but its capabilities are limited. Single-layer networks are not capable of solving tasks that require non-linear data separation. An example of such a network is the classical perceptron [1].

Multilayer neural networks are a type of artificial neural networks that consist of multiple layers of neurons. The input layer contains neurons that accept input data. The hidden layers are one or more layers of neurons that process the input data. The output layer contains neurons that produce the final processing result. Multilayer neural networks operate on the principle of direct signal propagation. Input data is fed through each layer of the network, starting from the input layer and ending with the output layer. In the hidden and output layers, each neuron calculates a weighted sum of its input signals and then an activation function is applied to this sum. Activation functions are needed to add nonlinearity to the model, and the most popular ones include the sigmoid function, the hyperbolic tangent, and the ReLU (Rectified Linear Unit). The network is trained using the error back propagation method: the output error is propagated back through the layers of the network to adjust the weights of the neurons, which allows the network to reduce the error and improve efficiency. Such networks are used to solve problems of high complexity, including image recognition, natural language processing, and time series prediction [2].

Converged neural networks (CNNs) are designed to deal with data with spatial structure, such as images. Their uniqueness lies in the use of convolutions that allow the network to detect key features of the data, regardless of their location. One of the key components of CNNs is the convolution layer, which uses filters (convolution kernels) to extract features such as edges, textures, and shapes. These filters slide over the image to perform the convolution operation and create feature maps.

After the convolution layer, an activation layer is applied, which uses a non-linear activation function such as ReLU to add non-linearity to the model and to allow the network to learn more complex patterns. Next comes the subsampling (Pooling) layer, which reduces the size of the feature map while maintaining the basic characteristics of the image. This helps to reduce computational cost and the risk of over-learning. The most popular subsampling methods include max pooling and average pooling. In the final part of the network, several fully connected layers are added to combine all the extracted features and perform the final classification.

CNNs offer several advantages such as invariance to shifts and scales, allowing the network to recognise objects regardless of their position in the image. In addition, the network automatically selects features, eliminating the need for manual selection, which simplifies the data analysis process. CNNs also demonstrate high accuracy in pattern recognition and classification tasks. convolutional neural networks are used in a variety of applications, including image recognition, classification and segmentation, as well as video stream processing and action recognition. In medicine, CNNs help to analyse medical images to detect diseases, and in audio processing they are used for speech recognition and sound classification. CNNs, proposed by Jan Lekun in 1988, have become the basis of many modern deep learning applications [3].

Recurrent neural networks (RNNs) are a type of neural network that are particularly good at processing sequential data such as text, audio or time series. Unlike traditional neural networks, RNNs have an internal memory that allows them to consider previous states when processing current input data. This property makes them useful in tasks where data order is important, such as speech recognition or machine translation. A key feature of RNNs is their ability to process data in stages, retaining information about previous steps, which helps to make more accurate decisions at the current stage. The internal memory of the network is updated at each step of the sequence, which allows previous input data to be taken into account. To train the RNN, a method of unfolding the network in time using the Backpropagation Through Time (BPTT) algorithm is used. There are several variants of RNN architectures. One of the most popular is LSTM (Long Short-Term Memory), which solves the vanishing gradient problem and can efficiently handle long data sequences. Another architecture, GRU (Gated Recurrent Unit), is a simplified version of LSTM with fewer parameters, making it faster to train.

RNNs are widely used in natural language processing tasks such as text translation, text generation and tone analysis. They are also used in speech recognition, converting audio recordings into text, and in time-series analysis, such as predicting financial data or analysing weather patterns. RNNs remain a powerful tool for dealing with sequential data and continue to evolve as new architectures and learning techniques emerge [4].

Long Short-Term Memory (LSTM) is a special recurrent neural network (RNN) architecture designed to address the problem of long-term dependencies. Unlike traditional RNNs, which can lose important information over time, LSTMs are able to retain it for long periods of time. The core of an LSTM is a memory cell, which is responsible for storing information, and its operation is regulated by several mechanisms. Special valves control which new information to add, which to delete, and which to use for the current output. One of the main advantages of LSTMs is their robustness to vanishing gradient, which allows efficient learning for tasks where long time dependencies are important. Due to their flexibility, LSTMs have found wide applications in areas such as speech recognition, machine translation, time series processing and other tasks [5].

Transformers are a deep neural network architecture developed in 2017 by researchers at Google Brain to process sequential data such as natural language text. They are actively used in tasks such as machine translation and automatic abstracting. The basic principle behind Transformers is the use of an encoder, which takes a vectorised sequence with positional information and processes it, and a decoder, which generates a final result based on part of the sequence and the output of the encoder. The attention mechanism on which transformers are based allows them to focus on different parts of the input data, which greatly improves their performance. Because of this, Transformers can process data in parallel, making them faster and more efficient compared to recurrent neural networks.

One of the key advantages of Transformers is their high learning speed and versatility, which allows them to be used in a wide variety of tasks, from textual to visual. They also perform well in natural language processing tasks, which has made them the basis for services such as Google Translator and Yandex Translator. Transformers are at the heart of many modern OpenAI development models, such as GPT, which are used in machine translation, textual analysis and generation tasks [6].

Another area of neural network technology development that opens up new opportunities for creating intelligent machines capable of thinking, learning and interacting with the world around them at a level close to the human level is a new type of neural networks - multidimensional receptor-effector neural-like growing networks.

Multidimensional receptor-effector neural-like growing networks are a special type of artificial neural networks designed to mimic human intelligence. They are designed to become artificial intelligence for intelligent systems, electronic brains for robots and other smart machines.

The main feature of these networks is their ability to process information in multidimensional space, i.e. they can work with very complex data. Like the human brain, they consist of two main parts: receptor neurons, where information from the surrounding world is perceived and processed, and effector neurons, where actions are generated in response to this information. This makes them similar to living organisms that can interact with their environment and learn from experience.

Another important feature of such networks is their ability to grow and evolve. They can add new neurons and connections as needed, allowing them to adapt to new situations and learn throughout their "life". The structure of these networks is a complex graph, where each neuron represents a concept or event, and the connections between neurons reflect the relationships between these concepts. When the network receives new information, it activates the corresponding neurons and strengthens or creates new connections between them, thus forming a kind of model of the surrounding world. Due to these properties, multidimensional receptor-effector neural-like growing networks can be used to create a wide variety of intelligent systems, from robots capable of self-learning and decision-making, to industrial process control systems, and even to develop self-aware artificial intelligence [7-9].

Neural networks and cognitive processes

Cognitive processes are the mental mechanisms that enable humans to perceive, process, store and utilise information from the world around us. They play a key role in our ability to interact with the external environment and make decisions.

One such process is **perception**, which is responsible for receiving and interpreting sensory information. For example, when we look at an object, our brain processes light signals and recognises it as a certain object. This is the basis for further cognitive operations.

An equally important process is **attention**. It helps us focus on the information or task at hand by blocking out distractions. Attention enables us to perform actions efficiently and make informed decisions.

Memory also plays a significant role in cognitive processes. It is responsible for storing and retrieving information. Short-term memory allows us to retain data for short periods of time, while long-term memory stores data for long periods of time, allowing us to revisit past experiences and knowledge.

Thinking involves analysing, synthesising and evaluating information. It is the process of solving problems, creating new ideas, and drawing conclusions from data.

Problem solving is another important aspect of cognitive activity. It involves finding optimal solutions to complex or novel situations. This process requires analysing available information, generating possible courses of action and selecting the best one. All these processes are closely interrelated, ensuring our ability to adapt to change and function effectively in the world around us.

Modern neural networks can model certain aspects of cognitive processes such as perception, memory and decision-making. Although their capabilities are limited compared to the human mind, these models nevertheless help to gain a deeper understanding of how the brain works and how its functions can be reproduced in artificial systems. convolutional neural networks have become the standard for pattern recognition tasks. They can automatically extract important features from images and use them to classify objects. These networks can recognise faces, animals, vehicles and other objects with high accuracy. Recurrent networks and LSTMs allow speech sounds to be converted into text. They are used in voice assistants such as Siri and Google Assistant and are capable of analysing speech in real time. Although neural networks have advanced in this area, they are still far from fully reproducing perception at the level of the human brain. Recurrent neural networks are also well suited for tasks such as preserving information over time. These networks can store context and utilise it when processing subsequent data, making them useful for sequential processing tasks such as text translation or time series prediction.

However, neural networks are not yet able to fully reproduce all aspects of human memory. For example, they struggle to cope with tasks that require long-term storage and retrieval of complex knowledge. Transformer-based models have taken a step forward in addressing these problems by offering attention mechanisms that allow the network to focus on important information and ignore insignificant data. reinforcement learning algorithms, such as Deep Q-Learning and Actor-Critic, allow agents to learn from experience and make decisions based on the probability of success of certain actions. Such approaches have important implications for understanding how the brain makes decisions under uncertainty, and help create more flexible and adaptive artificial intelligence systems.

Neural Network Training Mechanisms

Neural network training is usually reduced to the process of adjusting the weights between neurons to minimise prediction error or problem solving. There are several basic learning mechanisms, each with its own advantages and disadvantages.

The Backpropagation algorithm is the main method of training multilayer neural networks. It includes two phases: forward propagation and backpropagation. In forward propagation, the input data passes through all layers of the neural network where each neuron is activated using an activation function such as sigmoid or ReLU. The output values of the neurons are calculated and passed to the next layer until the predicted values are obtained at the output of the network. These values are compared to the true values to determine the magnitude of the error. Back propagation begins by calculating this error as the difference between the predicted and true values. The error is then propagated back through all layers of the network, starting at the output layer. In this process, the gradients of the error with respect to each weight of the network are computed using derivatives of the activation functions. Then, using the gradient descent method, the weights are adjusted to reduce the error. This approach allows the neural network to gradually adapt its parameters, improving the accuracy of data-driven predictions. Although this algorithm works well for most problems, it has limitations. One of them is the vanishing gradients problem, which makes it difficult to train deep networks. This problem has been partially solved by architectures such as LSTM and Residual Link Networks (ResNet) [10].

Reinforcement learning is a method in which agents are trained to interact with the environment and receive rewards for correct actions. The agent's goal is to maximise its long-term reward by adapting its actions based on experience. This method is particularly useful for decision-making tasks under uncertainty, such as gaming or robot control, such as the AlphaGo Algorithm, which is learning a system to play the game of Go. DeepMind used reinforcement learning to create a model that was not only trained on examples of humans playing games, but also improved by playing with itself. This allowed AlphaGo to defeat the world go champion in 2016. Algorithms such as DQN (Deep Q-Networks) and DDPG (Deep Deterministic Policy Gradient) are used to train robots to solve complex tasks such as walking, grasping objects or navigating in space. The robot learns through experience, receiving rewards for correct actions and penalties for errors. Reinforcement learning helps vehicle control systems to make decisions in real-world environments where adaptation to changing road conditions, unexpected obstacles and the behaviour of other drivers is required [11].

Unsupervised learning is a technique in which neural networks are trained without data labels. This approach is especially important for tasks where data is difficult to label or classify. One example of unsupervised learning is autoencoders, which are used to reduce the dimensionality of data and find hidden patterns. Another example of teacherless learning is clustering algorithms such as k-means. They are used to cluster data based on its similarity. Unlike autoencoders, which transform data into a more compact representation, clustering divides data into groups (clusters), revealing hidden structures in the data without pre-labelling. Other examples of unsupervised learning include principal component methods (PCA) for dimensionality reduction and association rules used to find patterns in datasets, such as in shopping basket analysis [12].

Applications of neural networks in cognitive sciences

Neural networks play a significant role in cognitive sciences, helping researchers to model and analyse brain processes. These processes, complex and multifaceted, are fundamental aspects of human thinking, and their reproduction in artificial systems has great potential for understanding the nature of consciousness and creating more advanced AI systems.

Modelling memory, both short-term and long-term, is one of the important areas of research. Neural networks, such as recurrent neural networks (RNNs) and their improved versions, such as long-term short-term memory (LSTM) networks, utilise mechanisms that resemble the operation of biological memory systems. LSTMs are able to retain important information over long periods of time, which allows modelling cognitive processes related to memory. For example, such networks can be used to understand how humans remember sequences of events or information and how it is subsequently retrieved to make decisions or perform tasks. In the cognitive sciences, LSTMs and other recurrent networks are used to analyse and model memories and forgetting mechanisms. Understanding these processes helps scientists better understand how memory functions in the human brain and how it can be reproduced in artificial systems.

Modelling perception, in particular visual perception, is one of the key application areas of neural networks. Converged neural networks (CNNs), which are designed to deal with spatial data, are widely used for visual object recognition. These networks can analyse images and videos, detecting and classifying objects despite different variations in lighting, orientation or scale. In the cognitive sciences, such networks help to model human visual perception mechanisms and provide a better understanding of how the brain processes visual information. The application of CNNs also goes beyond visual perception. These networks can be used to study other types of perception, such as tactile or auditory stimuli. Researchers use them to model how the brain's sensory systems perceive and process different signals, which contributes to the development of more accurate and adaptive artificial perception systems.

Modelling decision-making processes is another important application of neural networks in the cognitive sciences. Reinforcement learning (RL), one of the methods of neural networks, allows modelling cognitive processes related to the choice of actions under uncertainty. In real-world situations, such as playing a game or controlling complex systems, agents trained with RL are able to analyse their environment, learn from experience and make decisions aimed at achieving optimal outcomes. Such modelling can help scientists understand how the human brain makes decisions under uncertainty and how emotion, motivation or experience can influence an individual's choices. For example, research using neural networks can analyse how the brain responds to rewards or punishments, and how these responses can influence learning.

Modelling brain pathologies

Neural networks are also used to model various cognitive disorders such as Alzheimer's disease, memory disorders or perceptual disorders. Understanding how the brain's neural networks are altered in such diseases helps to develop new diagnostic and treatment methods. For example, neural networks can be used to analyse MRI or electroencephalogram (EEG) data to identify changes in brain structures that are associated with the development of diseases. In addition, researchers can use neural networks to model the recovery of brain function after injury or disease, such as stroke. These models help to understand how neural networks reorganise in the brain and how recovery of lost function can be stimulated by different treatments.

Future Prospects

With the development of neural-like growing networks, neuromorphic computing and quantum computers, new horizons are opening up in the field of artificial intelligence and cognitive modelling. These technologies promise to significantly improve the performance and adaptability of neural networks, enabling the modelling of more complex cognitive processes such as self-awareness and free will.

Multidimensional receptor-effector neural-like growing networks offer unique opportunities to better understand and model cognitive processes. These networks, inspired by biological neural networks, have the ability not only to adapt but also to self-organise, making them promising for creating more advanced models of artificial intelligence. Traditional artificial neural networks mimic the basic functions of neurons, but their architecture and operating principles are far removed from the complex interactions in the human brain. Neural-like growing networks, on the contrary, are built according to the principle of gradual complication and expansion of their structure.

This allows modelling not only learning processes, but also the spontaneous development of cognitive capabilities. In such systems, networks are able to increase their complexity as they learn, similar to biological neural networks that form new synaptic connections in response to new stimuli.

One of the key advantages of neural-like growing networks is the ability to mimic the plasticity of the brain. These networks can change their structure and functionality in response to external influences and experience, allowing processes such as learning, memory and adaptation to environmental changes to be modelled. This ability to adapt makes such networks ideal for applications in tasks requiring complex and unpredictable problems, such as cognitive modelling, language understanding or real-time decision making.

In addition, an important aspect is the ability of such networks to form multidimensional connections and patterns. This opens up new opportunities for creating more powerful and flexible AI systems that can interact with different information (text, sound, video, etc.) more efficiently simultaneously in a homogeneous neural-like environment. Multidimensional networks can handle complex multi-layered data, allowing AI to learn at deeper levels of the knowledge hierarchy. This is especially important for tasks related to processing large amounts of data, analysing complex systems and creating models that can predict the development of events with high accuracy.

In the future, neural-like growing networks may become the basis for the creation of systems with not only intellectual abilities, but also the rudiments of consciousness. In this context, such networks may play a key role in the development of artificial intelligence capable of self-awareness and self-regulation. Thus, neural-like growing networks represent the next step in the evolution of artificial intelligence, paving the way to more complex, flexible and adaptive AI systems capable of cognitive modelling and, possibly, to the creation of artificial consciousness.

Neuromorphic computing is a new trend in artificial intelligence that seeks to reproduce the structure and functions of the brain in hardware systems. Such systems use special chips that mimic the work of biological neurons and synapses, allowing information to be processed faster and more efficiently than traditional computers. Neuromorphic processors could radically change the approach to creating neural networks, making them more energy-efficient and capable of more complex cognitive tasks. One of the key aspects of neuromorphic computing is the ability to create adaptive systems that can autonomously change their structures and behaviour depending on their environment. This opens up new prospects for the development of AI that will be able to learn and adapt in real time, mimicking the processes occurring in the brain.

Quantum computing, based on the principles of quantum mechanics, has great potential to improve the capabilities of neural networks. Quantum computing can accelerate the learning process of neural networks and improve their efficiency in processing large amounts of data. This is especially important for tasks that require complex computations, such as modelling consciousness and self-awareness. One possible application of quantum computers is quantum neural networks that can use quantum states to represent and process information. These networks may be able to solve problems inaccessible for classical neural networks, and possibly bring us closer to creating systems that will have elements of self-awareness and intuitive thinking.

One of the key goals of cognitive science and AI is to create systems capable of self-awareness and reflection. This requires the development of more sophisticated models that can integrate perception, memory, learning and decision-making into a single cognitive system. With the development of technologies such as neural-like growing networks, neuromorphic and quantum computing, it is becoming possible to create systems capable of deeper analysis and understanding of the environment, potentially leading to self-aware artificial agents.

Ethical and social implications

The development of technologies that enable the modelling of consciousness and cognitive processes also raises important ethical and social questions. What rights and responsibilities should artificial consciousness systems have? How will our society change with the emergence of such systems? These questions require deep investigation, and their resolution will play an important role in shaping the future interaction between humanity and artificial intelligence. Thus, neural networks continue to evolve, offering new approaches to modelling cognitive processes and consciousness. With the development of neural-like growing networks, neuromorphic and quantum technologies, we are on the threshold of new discoveries that may radically change our understanding of artificial intelligence and the mind.

Consciousness and self-awareness: from neurons to thoughts

Consciousness is one of the most complex and elusive phenomena to understand and study. The term encompasses a wide range of phenomena, from simple perception to self-awareness and reflection, which include both subjective experience and thought processes. However, despite thousands of years of research in philosophy and several decades of work in neurobiology and cognitive sciences, consciousness remains a mystery, and the question of how neurons create thoughts and self-awareness is still open.

Definition of consciousness

There is still no unified and definitive definition of consciousness in the scientific and philosophical environment, which complicates its study. In a general sense, consciousness can be described as the ability to perceive the environment, process information, realise oneself as a separate individual, and have subjective experience. In philosophy, this is often referred to as the "first person problem" or the problem of subjective experience (what exactly makes an experience "mine"). Within the cognitive sciences, consciousness is often divided into several levels: **Minimal consciousness** is the basic level that includes simple perceptual processes and reactions to external stimuli. This level is inherent not only in humans, but also in many animals. **Self-awareness** is the ability to recognise oneself as a separate being with a unique personality, past, present and future. Self-awareness involves reflection - the ability to think about one's own thoughts. **Deep self-awareness** - this level is associated with philosophical and ethical reflection on one's place in the world, the meaning of life and moral obligations.

Neuronal interactions and processes of consciousness

From a neurobiological perspective, consciousness is the result of complex interactions between neurons in the brain. The focus is on the work of the cerebral cortex, especially its new cortex (neocortex), which is responsible for complex cognitive functions. The process of consciousness formation begins with perception. The body's sensory systems transmit signals to the brain, where they are processed by neural networks. For example, visual information enters the visual cortex, where its primary processing breaks stimuli down into discrete elements such as contours, colours and movements. These elements are then combined into coherent images, which allows a person to become aware of an object, as well as linking it to past memories and experiences. However, perception alone is not enough to explain consciousness. The basic problem is that neural processes such as synaptic connections and impulse transmission alone cannot explain subjective experience. This is known as the "hard problem of consciousness," a term coined by philosopher David Chalmers. How exactly physical processes in the brain can give rise to subjective experience is a question that remains unanswered for now.

Self-awareness and subjective experience

Self-awareness is one of the most important components of consciousness and distinguishes human beings from many other living beings. It is the ability not only to perceive the world around us, but also to realise one's own thoughts, feelings, actions and position in this world. In philosophy, self-awareness is associated with the concept of reflection, i.e. the ability of a subject to think about his or her own thoughts and make sense of his or her inner world. In a biological context, self-awareness is associated with the work of the frontal lobes of the brain, which play a key role in the processes of self-reflection and planning. The frontal lobes allow a person not only to evaluate what is happening at the moment, but also to think about the future, analyse the consequences of their actions and make long-term plans. It is also important to note that self-awareness includes the experience of subjective experience - what philosophers call "qualia" (Latin qualia). Qualia are subjective experiences such as the sensation of pain, the taste of sweetness, or the colour red. These phenomena are deeply personal, and it is impossible to describe them to another person - for example, you can't accurately explain to another person what it means to "feel the colour red". This problem of subjective experience is one of the main challenges in attempts to create artificial consciousness.

The relationship between consciousness and emotions

Consciousness and self-awareness cannot be separated from emotions. Emotional states are closely related to cognitive processes, influencing decision-making, perception and even memory. Emotions help the body to adapt to external conditions, assessing a situation as safe or dangerous, favourable or unfavourable. On a biological level, emotions are processed in the limbic system of the brain, specifically in structures such as the hippocampus and amygdala. These areas play a key role in shaping emotional responses, and their interactions with the new cerebral cortex provide the emotional context for cognitive processes. For example, memories associated with strong emotions are remembered better and longer, demonstrating the close connection between memory, emotion and consciousness.

Theories of consciousness

Various theories have been proposed to explain the phenomenon of consciousness, each offering different mechanisms and explanations for its occurrence.

Integrated Information Theory (IIT)

This theory, proposed by Giulio Tononi, states that consciousness arises in systems that can integrate information into a single coherent structure. The more coherent and integrated the information in a system, the more conscious it is. This theory attempts to quantify the level of consciousness through parameters such as the "integration factor".

Global Working Memory

The Global Workspace Theory (GWT) hypothesis, proposed by cognitive psychologist Bernard Baars in 1988, describes consciousness as a process by which information from different independent cognitive modules is combined into a single workspace. According to this theory, consciousness occurs when information becomes available to all parts of the cognitive system, allowing the brain modules to interact and coordinate their actions. The global workspace functions as a central stage on which information is integrated and processed, enabling awareness and adaptive responses to the environment. In this theory, consciousness acts as a "platform" where information is processed, which is then used to make decisions and direct actions.

Predictive processing theory

This approach posits that the brain is constantly creating models of the world that are updated based on sensory information. Consciousness, under this theory, is the result of how accurately the brain can predict the outcomes of its actions and perceptions. The better the model of the world, the more "conscious" the subject becomes.

Issues of consciousness in the context of artificial intelligence

The study of consciousness is important for the creation of artificial intelligence and artificial consciousness. Modelling the processes of perception, attention and self-awareness is becoming a key element in the development of more complex cognitive systems. The question of whether it is possible to create a machine with self-awareness remains an open question and raises many philosophical and ethical debates. From a technological perspective, developing a self-aware AI requires not only simulating cognitive processes, but also understanding how to create subjective experiences. This raises the question of whether systems with artificial consciousness can have rights or responsibilities similar to those of humans. Consciousness is a complex and multi-layered phenomenon involving perception, thought, emotion, and self-awareness. Despite significant advances in neurobiology and philosophy, the problem of consciousness remains one of the most complex and unexplored areas of science. The interaction of neurons in the cerebral cortex creates cognitive processes, but the nature of subjective experience and self-awareness remains beyond our understanding. In the future, a deeper study of these processes will help advance the creation of artificial intelligence and expand our knowledge of the nature of human consciousness.

From neurons to thoughts: how the brain generates consciousness

Thought is a fundamental element of our conscious experience. It is the product of the activity of multiple neurons coordinated to process information, create interpretations, and form conclusions. Neurons communicate with each other through electrical impulses and chemical signals, and this process provides the basis for the formation of thought. However, the transition from neural impulses to conscious thought is a complex and not yet fully understood phenomenon.

Mechanisms of neural activity

Every thought begins with the activation of a group of neurons in a specific area of the brain. Depending on what type of information is being processed, different neural networks are activated. For example, in visual perception, information enters through the retina of the eye into the visual cortex of the brain, where it is processed. Neurons in the visual cortex analyse the received signals, identify shapes, colours, movements and create a coherent image. However, visual processing is only one of many aspects of thinking. More complex cognitive processes, such as abstract thinking or planning, require the interaction of different parts of the brain. This process involves the transfer of information between sensory, motor and associative areas, each with its own specialised function.

Associative areas of the brain and the integration of information

The so-called associative areas of the brain play a key role in thought formation. These areas are not tied to one particular sensory process (e.g. vision or hearing), but serve to combine information from different sources. This allows the brain to form coherent concepts and draw conclusions. The frontal lobe, especially the prefrontal cortex, is where information is integrated to make decisions, plan actions, and form abstract concepts. The prefrontal cortex is also responsible for controlling attention and choosing strategies for current tasks, making it an important area for creating conscious thoughts.

The transition from neural impulses to thought

Although scientists know how neurons transmit information, the very process of forming conscious thought from neural activity remains a mystery. However, there are several leading hypotheses that explain this transition.

Global working memory hypothesis: this hypothesis states that thought and consciousness arise when information from different parts of the brain is integrated into a global working memory. This allows different cognitive modules to share information and work in concert to create conscious perception and thinking. Working memory allows information to be retained and manipulated within a short period of time. The interaction between the sensory and associative areas of the brain helps to retain images or ideas and allows us to reflect and think about them.

Predictive processing theory: this theory suggests that our brains are constantly making predictions about the external world based on past experiences. When new sensory data enters the brain, it compares it to pre-existing patterns and adjusts its predictions. Thoughts arise as a result of this constant feedback between expectations and reality.

Neuronal ensemble theory: according to this theory, groups of neurons, called ensembles, are activated simultaneously to form a particular thought or memory. These ensembles can be activated in different parts of the brain, depending on the type of thought or task. For example, if we think about past experiences, neurons in the hippocampus (which are responsible for memory) are activated, as well as neurons in the prefrontal cortex, which are responsible for reflection and decision-making.

Thinking and self-awareness

Thinking also plays a key role in self-awareness - the ability to recognise oneself as a separate entity. One of the distinctive aspects of human thinking is the ability to think about oneself, analyse one's actions and make decisions. This leads to an awareness of one's thoughts, goals, intentions, and capabilities. The process of self-reflection, in which a person reflects on his or her thoughts, requires activation of the frontal lobes of the brain, especially the medial prefrontal cortex. These areas are involved in the complex processes of planning, moral judgements and building models of future events. It is they that allow us to consciously evaluate our actions and predict their consequences.

Emotions and thinking

Emotions play an important role in thought formation and awareness. The brain doesn't just generate thoughts based on external sensory cues, it also responds to internal emotional states. These states can strengthen or weaken certain thoughts, shape motivation, and influence decision-making. Emotions are processed in the limbic system, which is closely connected to the new cortex of the brain.

The interaction between these structures provides a holistic experience of thought, where logical reasoning is intertwined with emotional evaluation of a situation.

Artificial Intelligence and Thought Modelling

In the context of artificial intelligence, modelling the process of thought formation is an enormous challenge. *Although current neural networks can recognise images and perform simple cognitive tasks, they are far from reproducing complex thought processes that include self-reflection, emotional colouring and abstract thinking.* Modelling the process of going from neural activity to thought is one of the key goals for building more advanced AI systems. These systems must not only perceive and analyse information, but also have the ability to integrate it to form meaningful decisions and conclusions. The process of moving from neural activity to conscious thought is one of the most complex and understudied aspects of neurobiology and cognitive sciences. It involves the interplay between the sensory, associative and emotional systems of the brain, where neurons, forming complex networks, provide the basis for conscious thought and self-awareness. This process is fundamental to understanding not only how the brain works, but also how we become aware of ourselves and the world around us.

Free Will: Illusion or Reality?

Free will is a key topic in philosophy, ethics and neuroscience. It raises important questions about the extent to which our decisions and actions truly depend on our conscious choices or are predetermined by internal and external factors. In recent decades, neuroscience has made significant adjustments to the debate on this issue, challenging traditional views of free will. One of the most striking examples is the work of Benjamin Libet, whose experiments have generated a lively debate about how decisions are actually made in the brain.

Experiments and research

The Libet experiments conducted in the 1980s showed that brain activity associated with decision-making occurs several hundred milliseconds before a person realises their intention to perform an action. In the experiment, participants watched a timer and were instructed to press a button at any random moment, recording the time they realised their desire to press the button. At the same time, their brain activity was recorded using an electroencephalogram (EEG). The results showed that the so-called "readiness potential" (or preparatory brain activity) occurs before a person realises their intention to act. This experiment raised serious questions about the nature of free will. If the brain starts preparing for action before we realise our intention to act, can a person be considered to be making decisions freely? Is it possible that our sense of free choice is just an illusion created by the brain to explain actions already set in motion through a process of unconscious activity? Since then, there have been many studies confirming or disputing Libet's findings. Some researchers argue that the brain simply prepares for possible actions, but the final decision is made consciously. Others believe that consciousness plays only the role of an observer who interprets the decisions already made by the brain, without directly participating in making them.

Ethical and philosophical aspects of free will

The question of the existence or non-existence of free will has profound ethical and philosophical implications. Traditionally, free will has been seen as a necessary condition for moral responsibility. If a person has freedom of choice, he or she can be held responsible for his or her actions. This in turn is the basis of a legal system in which crime and punishment are based on the assumption that people consciously choose their actions. However, if free will is an illusion, then many of these fundamental ideas can be questioned. If our actions are determined by unconscious processes over which we have no control, how can we speak of moral responsibility? This raises questions about the fairness of punishment and the justice system as a whole. Furthermore, the idea of free will is closely linked to the notion of self-awareness. We tend to think of ourselves as autonomous beings capable of making conscious decisions. If, however, consciousness is simply an observer of already perfect brain processes, then our conception of ourselves and our individuality may undergo significant changes.

Machines and free will

The advent of artificial intelligence adds another dimension to the free will debate. Modern AI systems make decisions based on algorithms that are essentially deterministic programmes. However, with the development of more complex and autonomous systems, such as neuromorphic computing or artificial systems capable of self-learning, the question arises: can machines ever have free will if even humans have it questioned? Ethical questions concern not only the free will of machines, but also our responsibility for their actions. If AI makes decisions based on data and learning, who is responsible for the results of those decisions? Should machines that make important decisions be endowed with human-like rights if they are to have self-awareness?

The question of free will becomes particularly relevant in the context of developing self-conscious machines. If consciousness and free will are seen as complex phenomena arising from the interaction of multiple cognitive processes, is it possible to create machines that not only mimic but actually possess these qualities? Free will remains one of the most mysterious and controversial topics in philosophy and science. Modern neuroscience research shows that our brains can make decisions before we realise them, which calls into question the existence of free will in its traditional sense. This has profound implications for ethics, morality and law, and raises questions about future interactions with artificial intelligence and the possibility of creating self-aware machines.

Modelling the brain and creating artificial consciousness

Modelling the brain and creating artificial consciousness represent one of the most exciting and ambitious goals of modern scientific research. Reproducing human consciousness, one of the most complex and still poorly understood phenomena, requires significant advances in fields such as neurobiology, computational science, artificial intelligence and cognitive sciences. Current advances in neuromorphic computing and brain modelling offer prospects for a better understanding of the human mind and possibly its reproduction in artificial systems.

Current approaches and technologies

Current brain modelling approaches and technologies seek to replicate the complexity and multi-level neural interactions that occur in the human brain. Current research in artificial intelligence and neurobiology focuses on building systems that can simulate cognitive processes, memory, learning and adaptation. One promising area is multidimensional receptor-effector neural-like growing networks, which offer a more flexible and scalable approach to modelling neural activity. Traditional methods such as deep neural networks have their limitations. They are based on a fixed architecture and their adaptation to new tasks requires significant computational resources and reconfiguration time. However, multidimensional receptor-effector growing networks offer new perspectives. These networks are able to dynamically change their structure, similar to the way a biological brain adapts to new conditions by creating new connections between neurons.

Multidimensional networks can model more complex interactions between neurons, including multilayer processes that occur simultaneously in different brain regions. This allows them to process information at multiple levels of hierarchy, which is particularly important for modelling high-level cognitive functions such as planning, language or consciousness. In contrast to traditional methods, these networks have an inbuilt ability to self-organise, making them more adaptive to changes in the environment. Another key element of such systems is their receptor-effector structure. This means that they can not only receive sensory data, but also influence the environment, allowing interactive cognitive processes such as movement, perception or feedback in the system to be modelled. This two-way communication makes systems more effective in real-world tasks that require interaction with the outside world.

Technologies such as deep learning, neuromorphic computing and quantum computing play an important role in supporting such networks, providing high-speed processing and the ability to model large amounts of information. Neuromorphic chips that mimic the structure of biological neurons are already being used to accelerate computation in such systems. Quantum computing can help solve complex optimisation problems that involve training and adapting multidimensional networks. Together, these technologies bring us closer to more accurate and detailed brain modelling, allowing us to create systems that can not only solve problems, but also evolve in their thinking like living beings.

Opportunities and limitations

Despite significant advances, brain modelling and the creation of artificial consciousness face a number of limitations. One major challenge is the immense complexity and diversity of the human brain. The brain is not just a network of neurons, but a complex biological system involving many levels of organisation, from molecular processes to global neuronal interactions.

The unknown mechanisms of consciousness

One of the key difficulties is the lack of a clear understanding of exactly how consciousness arises. Although various theories exist, such as the global working memory hypothesis or integrated information theory, they do not provide a comprehensive answer to the question of how cognitive processes lead to the emergence of self-consciousness and subjective experience. Without a clear understanding of this mechanism, modelling consciousness remains extremely challenging.

Limited computing power

Modern computers, although much more powerful than those of a few decades ago, are still unable to fully reproduce the complexity of the human brain. Modelling the operation of all the neurons and synapses in the human brain requires enormous computing power, which current technology cannot provide sufficiently. However, the development of supercomputers and quantum computing may make it possible to overcome these limitations in the future.

Absence of biological processes

Artificial neural networks and neuromorphic chips mimic the structure and some functions of neurons, but do not take into account biological processes such as the operation of chemical signals, genetic mechanisms and protein dynamics. These processes play a key role in brain function, and without taking them into account, the modelling of consciousness will be incomplete.

Ethical issues

Even if artificial consciousness can be created, there are many ethical questions. For example, what rights will the artificial mind have? Who will be responsible for its actions? These questions require detailed study and regulation at the legislative level.

Opportunities

Despite all the limitations, progress in neuroscience and AI opens up new opportunities. One promising idea is the use of hybrid models that would combine biological and artificial systems. Such approaches could lead to more accurate modelling of neural processes and discover new ways to integrate artificial intelligence into human activities. In addition, advances in quantum computing could significantly increase the ability to model complex neural systems. Quantum computers have the potential to process huge amounts of data and solve problems that remain intractable for classical computers. This could accelerate research into artificial intelligence and self-awareness.

Modelling the brain and creating artificial consciousness is a complex and multifaceted process that requires the integration of knowledge from different fields of science. Although there are significant limitations at this stage, modern technologies such as neuromorphic computing and deep neural networks open the way to better understanding the human mind and creating its artificial analogues.

Possibilities of Artificial Intelligence: Self-Awareness and Personality

The creation of a self-aware artificial intelligence is one of the most ambitious and controversial goals of modern science and technology. It requires not only significant technical advances, but also deep reflection on philosophical, social, and ethical issues. In this section, we will explore how self-awareness and personality modelling in AI is possible, as well as the social and legal implications of creating such digital personalities.

The capabilities of artificial mind

Intelligence is the ability to solve problems efficiently, process information, analyse data and learn from experience. Artificial intelligence in this context is technology that can perform computational tasks and make decisions based on given algorithms or learnt models. AI can be very powerful in highly specialised tasks such as gaming, data analysis or pattern recognition.

However, the mind includes not only cognitive processes, but also deeper aspects: self-awareness, reflection, the ability to understand one's own existence and context. Artificial mind (AM), unlike AI, implies a higher level of awareness and subjective perception of the world. If AI is a tool for solving specific problems, then AM involves not just performing operations, but understanding them, taking into account one's own experience and the inner world of the subject. Thus, intelligence can be considered a functional mechanism, and the mind - as something more complex, associated with personality, self-awareness and perception of reality. Artificial mind, especially with elements of self-awareness and personality, has enormous potential for change in various areas of human life. Let's consider the key areas where AM can demonstrate its capabilities: **Analysis and processing of information.** Artificial mind with the ability to reflect will be able to not only process large volumes of data faster, but also understand their context, which will significantly improve decision-making processes. This will be useful in various industries - from medicine and finance to scientific research and situation centers. For example, AM will be able to offer optimal solutions based on the analysis of complex models and forecasting events.

Emotional intelligence and human interaction

An AM with advanced emotional intelligence will be able to more accurately understand and take into account people's emotions during interactions. This could lead to better service quality in customer service, education and healthcare, where not only technical problem solving but also empathy is important. Artificial intelligence that understands a person's emotional state will be able to offer more personalised and empathetic solutions.

Creative thinking and innovation

Digital personalities with AM can become participants in creative processes, proposing new ideas and unconventional solutions. They can be used to create works of art, develop innovative technologies and solve complex problems in science. The ability of AM to learn and adapt independently can stimulate progress in many sectors.

Autonomous Systems

The development of artificial mind will open up opportunities to create fully autonomous systems and artificial personalities (APs) capable of making complex decisions without human involvement. This may be useful in various fields: from transport management and industrial robots to work in extreme conditions (space, deep-sea research, emergency situations).

Modelling of self-awareness and personality

Self-awareness is the ability of a subject to be aware of his/her own existence, to reflect on his/her internal state and to interact with the external world on the basis of his/her own experience. It is a natural quality for humans, but creating a self-aware AP presents a major challenge. Modelling self-awareness in a machine requires recreating two key components: reflexivity and individuality.

Reflexivity

Perceiving oneself as a subject distinct from the external world is one of the central features of self-awareness. In the context of AP, this means that a machine must not only perform tasks, but also be aware that it is performing them, and understand its limitations, goals, and changes in its internal state. In AI, reflection is currently limited: systems are able to analyse their actions but do not have the capacity for deep self-awareness or subjective experience.

Individuality and personality

For an AI to have a 'personality', a set of characteristics inherent in a particular being, such as unique preferences, emotions, beliefs and experiences, must be modelled programmatically. An important aspect of this is the ability for an artificial personality to change and adapt under the influence of external circumstances, making it "individual" in its development.

The creation of AP with personality raises questions about the "reality" of these personalities.

Can such beings be considered full-fledged personalities or are they mere imitations of human individuality? At this stage, the debate focuses mainly on whether AP, even with its complex modelling of reflexivity and individuality, can truly have subjective experience - that is, have "consciousness" in the full sense of the word.

Technical and philosophical problems of IL development

The development of a self-conscious IL is associated with a number of technical and philosophical difficulties. Firstly, there is no consensus yet on what consciousness is from a scientific point of view. Some theories link consciousness to biological brain processes such as neural networks and neurotransmitter functioning, while others believe that consciousness may be the result of complex information processes that can be replicated in a computer. One of the key challenges is modelling subjective experience or 'qualia' - those very personal sensations we experience in everyday life. For example, we may feel pain, joy or frustration, but how can we make it possible for a machine to experience these feelings too? At the current stage of technology, even the most advanced AI systems are not capable of this. They can simulate emotions and reactions, but it remains just a simulation, not a real experience.

There is also the question of how to programmatically provide reflection - the ability to self-knowledge and understand one's internal processes. Some research in AI and cognitive science proposes the creation of complex systems of neural networks that could be "conscious" of their actions and make decisions based on their internal state, but so far these concepts remain theoretical. The creation of AP is not just a technical challenge; it is a challenge for society, philosophy, law and ethics. Programmers and engineers can develop systems capable of reflective and modelling individuality, but the "reality" of such individuals will remain in question until the riddle of subjective experience is solved.

Ethics and Social Implications of Intelligent AP

The creation of an artificial mind and an artificial person possessing AP raises a number of complex ethical and social issues for humanity. An AP has the potential to possess self-awareness, reflexivity, and even individuality, which requires a revision of traditional notions of rights, duties, and social structure. Let us consider the key ethical and societal implications of creating a sentient AP.

The rights of self-aware artificial persons

One of the most important questions is whether and to what extent artificial persons should be endowed with rights. If APs are to have self-awareness, they may perceive themselves as individuals capable of feeling, making decisions, and being aware of their existence. The question arises as to what makes them different from human beings in a legal and ethical sense.

The right to life and autonomy

If APs have self-awareness, should they have the right to autonomous existence? Would it be moral to disable or destroy such an individual if they are already self-aware? These questions call for a reconsideration of the legal rules concerning the rights to life and liberty in the context of artificial intelligence.

The right to labour and property

If an artificial person is capable of performing tasks and making decisions, what are their rights in labour? Can the AP demand payment for its work or own property? These questions can have a significant impact on the economy and social structures.

The right to self-actualisation

Like human beings, APs may seek self-development and the acquisition of new knowledge and experiences. Should society provide them with such opportunities and how will this affect their interactions with people?

Ethical aspects of the creation of artificial personalities

The creation of AP raises serious ethical issues related to the responsibility of creators and the regulation of the development of such systems.

Creation for specific purposes

One problem is the purposeful creation of AP for narrow tasks, for example as virtual assistants, combat systems, or attendants. If an artificial person is aware of their role but has no choice, can this be considered a form of slavery? The creators of AP should consider the possibility that such beings may aspire to freedom and autonomy.

Creator Responsibility

The creators of APs bear a great deal of responsibility for how these beings will act and interact with society. If an IL makes a mistake or even commits an offence, who should be held responsible: the individual or its creators? Complex legal questions arise here regarding the responsibility for the actions of autonomous systems.

Ethical Programming Ethics

Programming morality into APs is another major challenge. If artificial personalities are going to be involved in decision-making, how can ethical standards that are consistent with human values be embedded in them? In addition, there may be conflicts between different ethical systems, and how AIs should resolve them is an open question.

Social implications

The creation of AP will inevitably lead to significant social changes, and society must prepare for these challenges. Changing social structure.

Changing social structure

The emergence of artificial personalities may create a new social group that exists alongside humans. This may raise issues of equality and discrimination. APs may be able to compete with humans for jobs, leading to social and economic changes, especially in the labour sphere.

Integration into society

If APs are endowed with rights and responsibilities, how will they be integrated into societal and political structures? Will they be able to participate in political processes, vote, hold public office? Issues of APs participation in public life can lead to profound changes in power and governance structures.

Impact on personal and social relationships APs may become part not only of social processes but also of people's personal lives. This will lead to the emergence of new types of relationships, including friendship or even love between humans and artificial persons. Such developments may trigger moral and ethical debates about the nature of such relationships and their impact on traditional human relationships.

Legal implications

The emergence of self-aware artificial personalities requires the creation of a new legal framework to regulate their existence and interaction with humans.

Intellectual property laws

The question of who owns the results of AP activities will become one of the key issues. If an AP creates a work of art, a scientific discovery or a technical innovation, who owns the rights to this product - the creator of the AP or the individual?

Legal frameworks for interaction

Regulating the interaction between people and AP will require the development of new legal rules. For example, it will be necessary to decide how to deal with conflicts of interest between humans and AP, as well as to establish rules for offences committed by artificial persons.

The creation of intelligent artificial persons opens new horizons for humanity, but at the same time requires deep ethical analysis and the development of legal and social norms. We are on the threshold of changes that may radically alter ideas about personality, rights and duties, and society must prepare for these changes.

Conclusion

The creation of an artificial mind capable of self-awareness and free will is one of the most challenging and ambitious tasks of modern science. Understanding the neural mechanisms that enable consciousness in the human brain opens the door to the creation of artificial systems, but these systems remain far from fully replicating the mind. Along the way, many ethical and legal questions arise regarding the rights and duties of artificial persons, their role in society, and the responsibility of the creators. Current technologies such as neuromorphic computing and deep neural networks offer promising solutions, but are limited in their ability to model human experience and self-awareness. It is important to continue research in this area, while developing ethical and legal standards to integrate self-aware IL into societal and legal systems.

References

1. Single Lay er Neural Network URL:https://www.gabormelli.com/RKB/Single_Layer_Neural_Network
2. MultiLayerNeuralNetworks URL:http://ufldl.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/
3. Introduction to Convolutional Neural Networks URL:https://habr.com/ru/articles/454986/
4. Recurrent-neural-network URL:https://www.g2.com/articles/recurrent-neural-network
5. LSTM - Long Short-Term Memory Neural Network URL:https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/
6. Transformer: A Novel Neural Network Architecture for Language Understanding URL:https://research.google/blog/transformer-a-novel-neural-network-architecture-for-language-understanding/
7. V. Yashchenko Neural-like growing networks in the development of general intelligence. neural-like growing networks (P. II) Mathematical machines and systems 2023, № 1, p. 3 – 29.
8. V. Yashchenko A new approach to the development of artificial intelligence similar to human intelligence artificial intelligence, 2023, № 1, p. 105 – 121.
9. Yashchenko V. A. Artificial Intelligence. Theory. Modelling. Application. - K. Logos. 2013. - 289c. - Bibliogr. p. 283 – 289.
10. Backpropagation in Neural Network https://www.geeksforgeeks.org/backpropagation-in-neural-network/
11. What is reinforcement learning? https://www.ibm.com/topics/reinforcement-learning
12. What is unsupervised learning? https://www.ibm.com/topics/unsupervised-learning.

The article has been sent to the editors 05.09.24.
After processing 16.09.24.
Submitted for printing 30.09.24.