Epistemological Aspects of Edward Fredkin's Digital Cosmology in the Context of the Development of Artificial Intelligence Ideas

Andrii Synytsia

Doctor of Philosophical Sciences, Professor, Ivan Franko National University of Lviv
(Lviv, Ukraine)
E-mail: andrii.synytsia.edu@gmail.com
https://orcid.org/0000-0002-0983-7187

Synytsia, Andrii (2026) Epistemological Aspects of Edward Fredkin's Digital Cosmology in the Context of the Development of Artificial Intelligence Ideas. *Philosophy and Cosmology*, Volume 36, 193-205. https://doi.org/10.29202/phil-cosm/36/8

The article examines the main features of E. Fredkin's digital cosmology. Based on the analysis of the epistemological aspect of his concept, it is considered how he interprets the question of the possibilities of our knowledge of the world around us, defines the fundamental principle of the Universe, throught the analysis of the Other, and reflects on the world of human existence through the study of the soul in its static and dynamic manifestations. It is shown that despite the attempt to define the informational-computational approach as a universal one, in the study of the human inner world it appears to be imperfectly developed, and the further development of artificial intelligence theory has only confirmed that many assumptions about the digital nature of the world, which seemed obvious to Fredkin at the beginning for the transition from metaphors and parallels with the computer sphere to the construction of digital cosmology, as well as ontology and clarification of the metaphysical basis of reality, not only remained unconfirmed, but interpretations appeared that offered conclusions that differed significantly from them (while radicalizing certain ideas, replacing the real with the virtual, and developing not so much the idea of discreteness as the idea of continuity). In addition, attempting to explain how the world works by reducing everything to computational fundamentals provides a technocentric explanation of the nature of reality, which may be overly simplistic, as it leaves gaps in the explanations and does not solve the problem of consciousness and human experience. At the same time, it has been stated that the development of artificial intelligence technologies can clarify many issues of digital cosmology, in particular by improving our knowledge of the external world and human existence.

Keywords: digital cosmology, finite nature, information-computational approach, limits of cognition, reality, metaphysics of the Other, human inner world.

Received: 9 November 2025 / Accepted: 10 January 2026 / Published: 1 February 2026

© Synytsia, Andrii, 2026

Introduction

The idea of representing the foundations of our Universe based on the analysis of digital information processes gained popularity long before the advances in artificial intelligence became so significant that scholars even began to discuss the inevitability of technological singularity. At some point, it became evident that the digital approach, information-computational in its essence, was claiming the status of a universal one, capable of explaining not only the ideas of the macrocosm but also those of the microcosm. From a philosophical perspective, this approach advocates the idea of pancomputationalism, based on the assumptions of (i) finite nature, (ii) the existence of a first cause, and (iii) the interpretation of change as a process of computation.

It is not difficult to notice that these ideas are essentially not new to philosophy. Indeed, digital cosmology has developed and refined a number of philosophical ideas, including realistic, naturalistic, deterministic, reductionist, and monistic ones. Here it is worth mentioning the discreteness of matter according to the doctrine of atomism, as well as the idea of logical atomism, and ancient philosophical and theological explanations about the origin of the world, the necessity of the existence of some first cause, and the fact that the idea of describing changes in the surrounding reality as a process of implementing computation was formed entirely in the Pythagorean paradigm, according to which the world is fundamentally based on number and proportion, and for now we will not go into details about what these numbers are (discrete or continuous as series of real numbers), how many of them there are (two, three or more), and how the process of interaction between them is implemented. The fundamental feature of the digital philosophy approach is that it gives new meaning to all these philosophical ideas. Thus, not only material things, but also the attributes of matter – space and time – turn out to be discrete; the first cause begins to be interpreted from the standpoint of the computer paradigm as the Other, and any physical or mental relations as those that are conditioned by the process of computation and information processing at the most fundamental level of reality. And even though interpretation from the position of digital philosophy destroys traditional ideas about the continuity of physical phenomena, its proponents believe that this approach describes things and relationships as they are and does so much more accurately than mathematical abstractions that describe continuity using differential calculus and therefore have nothing to do with reality at all.

But to what extent is the information-computational approach justified not only in matters of interpretation of the cognitive sphere, but in ontology and cosmology in general? And with the success of artificial intelligence, has the information-computational approach to explaining the foundations of the Universe become more reliable? We will try to find answers to these questions by analyzing, in the perspective of the development of artificial intelligence ideas, E. Fredkin's digital cosmology, formulated primarily in his philosophical works "Finite Nature" (1992), "A New Cosmogony" (1993), "On the Soul" (2000), "An Introduction to Digital Philosophy" (2003) and supported by some research in physics. To achieve this purpose, we will focus on a critical analysis of Fredkin's ideas about the limits of our cognition, the foundations of the Universe, and human existence, and comparing them with later philosophical ideas formed in light of advances in artificial intelligence.

What can we know about the world?

The essence of the first question of digital cosmology, which needs to be clarified from a philosophical point of view, is what exactly we can understand about the foundations of

our world. Are we, with the capacities of our intellect, able to interpret correctly how the Universe is structured? Or perhaps it is more appropriate to take the position of an agnostic and maintain that such an endeavor is impossible – that it is like an attempt to examine an investigative instrument by means of that very instrument, or an attempt to understand issues for which we have neither the appropriate methods of knowledge nor a sufficient empirical basis. And this is despite the fact that the corresponding expressive capabilities of language and even of reason itself do not allow them to become representatives of certain complex ideas. Indeed, if we interpret the basic principles of reality only as something unspeakable, we will obtain a version of the worldview characteristic of Wittgenstein at the time of his "Tractatus Logico-Philosophicus", according to which, even though the limits of our world are outlined by the limits of language, there is something unspeakable (the world of ethics) that influences us and therefore we are able to comprehend it in a certain way, although we cannot express the results of our cognition by linguistic means (Wittgenstein, 1922, § 4.115). Much of what is not expressed in words can be imagined, experienced, or conceived, and for that very reason, it has a certain meaning. For his part, Fredkin prefers to think not so much about the nature of the unspeakable as about what is unthinkable (Fredkin, 2003: 241-242).

Unthinkable does not mean logically contradictory. The unthinkable may well be not just metaphysically possible, but even factually possible (if things like the finite nature hypothesis turn out to be correct). The fact that we can fully understand that there are things that we will never be able to understand is already a certain step in the direction of clarifying what is beyond our knowledge, that is, the unthinkable. It is important to understand here that logic, but not necessarily the one we express using formal means, goes beyond language. To a large extent, this is because the results of our cognition are recorded using a terminology system established in science. This very often determines our very process of describing things and phenomena of the objective world and the way we think about them, and therefore how we understand them. Thus, the unthinkable is such not only because of its complexity, but also because of language.

For Fredkin, it is quite obvious that the capacities of formal systems (mathematical frameworks) are not sufficient to describe the specifics of all the laws of the objective world. As is well known from Gödel's incompleteness theorems, formalization as such is limited in its expressive capacities. At the same time, the results of observations indicate the uncertainty present at the level of quantum mechanics, where the description of the seemingly simplest basic elements of reality becomes quite complex. However, as Fredkin argues, we should try to descend to an even more fundamental level of matter than elementary particles and explore the finite nature of space and time as basic attributes of matter, which will help us understand other important issues in modern theoretical physics, including those within the scope of quantum mechanics (Fredkin, 2004).

In general, it is appropriate to understand space and time as divided into cells, of course, extremely small, which cannot be confirmed experimentally at present. In any case, it is the idea of finitude that is important. As Fredkin explains: "Finite Nature means that the world is made up of digital information" (Fredkin, 1992: 349). In other words, the finitude of nature makes it possible to express the processes and phenomena of the objective world as informational, as well as to state that "…the amount of information in any small volume of space-time will be finite and equal to one of a small number of possibilities" (Fredkin, 1993: 116). These possibilities, which can be defined by two possible states of the bit (0 or 1) play the role of the fundamental alphabet of being, since their configuration determines the "computational" reality of being and makes it possible to express the laws of physics as

algorithms that process information and form the structure of space-time relations between things. It is important that "if finite nature is true, then there are certain consequences that are independent of the scale" (Fredkin, 1992: 345). And also that elementary units of time and space, like information, can be expressed only in integers, and, accordingly, it is integers that will be included in the formulas that characterize the enormous computational capacity of space-time, describe the laws of the objective world, and the nature of things.

According to the information-computational approach: "The meaning of the digits in Finite Nature is that the information process that the digits and cells are engaged in define the space and time and matter and energy of our world" (Fredkin, 1993: 118). In computer modeling, such discreteness of space-time structures is often used, but the available computing power of modern machines is, of course, not enough to model most physical processes. After all, to accurately model a system, the order of computational power of the models and systems would have to be comparable (Fredkin, 1993: 118). So it turns out that we have a theory, but it is unlikely that it will ever be possible to verify it. However, the lack of direct arguments that would prove the correctness of the assumption of finite nature, according to Fredkin's idea, should not be a reason for its rejection, given the development of science, in which it gradually became clear what role the same atoms or quanta of light (photons) play in the structure of matter and its manifestations. All this should lead us to think more about the finitude of nature in general than about its continuity. However, today we can also find alternative approaches that propose rethinking the very idea of continuity as a new methodological framework and demonstrate that, even at the level of finite things, there are continuities, as substantiated in Peirce's theory of synechism, which considers space, time, and laws as continuous (Pietarinen, 2025). Although we note that this alternative approach is valuable mainly in the process of researching issues of biology and bioengineering, which consider the issues of development and acquisition of new qualities. In many ways, the relevance of each approach is determined by the context of the research area. At the same time, it is also determined by the content of the formulations that we take as a basis in the process of conducting scientific research.

The discreteness of space and time, according to the requirements of digital philosophy, is a necessary condition for the transition from physical phenomena to computational processes that operate on information and are universal by their very nature, quite sufficient to explain various physical processes (Fredkin, 1990, 2005). It should be noted that Fredkin distinguishes only four laws of digital physics: "I. Information is conserved. II. The fundamental process of nature must be a computation-universal process. III. The state of any physical system must have a digital representation. IV. The only kind of change is that caused by a digital informational process" (Fredkin, 2003: 206). It is noteworthy that the foundations of nature, according to the approach of digital philosophy, cannot be non-computational. It is computational processes that make possible not only the construction of computers (these computing machines themselves indicate that the foundations of things are computational), but also life as such (we come to understand it gradually, making transitions from computational processes to physical, chemical, and finally biological ones). And although it is not easy to reach a correct digital representation characteristic of a particular state of nature, theoretically such transitions are quite unambiguous.

Even if this transition is accompanied by an explanatory gap, this does not mean that we cannot learn anything about the first cause of our world. It will be something like a retroanalysis task, when we have results and try to understand what the ultimate reality was that preceded them. Whether this understanding of ultimate reality will be close to the reasoning of E. Steinhart ("ultimate reality is a massively parallel computing machine sufficiently

universal for the realization of any physically possible world" (Steinhart, 1998: 117), or whether it will be a world of digital metaphysics of a completely different nature, as radically different from ours as the reality of software is different from the reality of hardware, or even more, can only be assumed.

For Fredkin's interpretation, it is important that the nature of computation itself be universal and remain so even in worlds that humans cannot think or imagine. Or in our world, if the study of its foundations required non-algorithmic understanding, a form of knowledge that goes beyond computation [Faizal, Krauss, Shabir, & Marino, 2025: 10]. Even though Fredkin himself does not develop the corresponding assumption, the opportunities that the development of artificial intelligence brings should not be neglected. It is quite likely that the findings of artificial intelligence and future advances in emerging technologies will contribute to a deepening of human understanding – initially through the attainment of new research results and the development of previously unknown methodological approaches, and later through their influence on human cognitive capacities themselves, in particular through genetic engineering, neuro-computer interaction, or even enhancing the potential of the collective mind. All this should expand the horizons of our knowledge and bring us closer to understanding the first causes of reality.

And all thanks to the flexibility of computation, the ability to use it to explain not only the nature of the surrounding world, but also the principles of cognition and the peculiarities of the functioning of the human mind. Chalmers draws attention to this kind of flexibility characteristic of computation in the process of explaining the nature of things when he argues that computation provides "a common framework within which many different theories can be expressed, and by providing a tool with which the theories' causal mechanisms can be instantiated" (Chalmers, 2011: 357). Computation itself proves useful not only for explaining the nature of physical reality, but also what lies beyond it. This is undoubtedly its advantage; however, as will be demonstrated below, in the quest to go beyond the limits of our reality or in the study of the nature of human existence, things are far from unambiguous.

Knowing the nature of the Other

The question is, how exactly is it possible to apply a computational model to explain the transcendent? The fact is that the finite nature is indicated not only by discoveries such as the atomic structure of matter, but also by the establishment of general principles of cellular automata. Knowledge of these principles of operation suggests the need to find out more about the foundations of the world. Cellular automata are actually specialized computers capable of functioning independently according to given algorithms, ensuring the transition of cells from one state to another, which actually turns the system into a self-organizing one. The very fact that nature is finite makes it possible to model all processes using a universal machine, some cellular automaton that operates on bits of information (Fredkin, 1990: 259). It should be noted that information expressed in bits, as well as space and time, are discrete units for measuring nature. The meaning of information, which is necessary for explaining nature, can only be understood within the framework of a certain information process. However, some processes are so complex that, despite their determinacy, the resources for their computation must be infinite.

The need for unlimited resources, like the nature of computation itself, raises the question of why things are the way they are and not otherwise. According to Fredkin's idea, we may not understand many things and not know how they work, but it is important to understand the general logic of fundamental processes. In other words, to understand how computation

works: "What is certain is that worlds that are qualitatively beyond our power to imagine are also capable of supporting computation" (Fredkin, 1993: 119). Knowing this fact is enough to ask where the ultimate computer is. However, the answer Fredkin himself provides seems unlikely to satisfy many. After all, by asserting that the ultimate computer is located elsewhere, Fredkin acknowledges the absolute difference between this other place and everything we are accustomed to dealing with. It can be reasonably assumed to be a place with a structure that we do not understand, and with laws that we cannot describe using the logic and terminology we are accustomed to. To at least somehow appreciate the Other, we must first learn a great many things about its creation – about our own world. In particular, to know the size of our Universe and whether it will expand infinitely and whether this expansion will ever stop, as well as to establish what excess computational power our Universe possesses. This would make it possible to estimate the computational resources invested in the functioning of our Universe, at least somehow narrowing down the number of reasons for its existence. Although it is possible that knowledge of these facts may remain beyond our understanding only due to the limited capabilities of our mind, which will always lack empirical facts, or sensory data, which, given the scale of the issues under study, we will simply never be able to achieve.

Continuing his reasoning (largely in terms of Aristotle's doctrine of four causes), Fredkin reflects not only on form, matter, and efficient cause, but also on the purpose of the universe's existence. However, this purpose is the result of all the computing resources of the universe and will remain incomprehensible. Although Fredkin says nothing about the possibility of studying the direction of development of the Universe, does not cover all possible computational resources, but resorts to the analysis of particular cases and further prediction, as well as to the extrapolation of knowledge from one subject area to another. Under such conditions, there will always be a possibility that at some point the interpretation of the meanings of individual processes may change beyond recognition (as in Nelson Goodman's "grue" paradox), but at least such a prediction makes it possible to understand this purpose, starting from the information resources accessible to humans. Ultimately, one might wonder whether this Universe is a final creation, or simply part of more complex and fundamental computational processes that operate at higher levels of reality with an abstract logical-mathematical and informational structure and its own physical laws, different from ours. It is quite likely that such primordial universes could be the first cause of the emergence of universes alternative to ours. The identification of the structure of such alternative Universes (possible worlds) would allow us to further clarify what the Other is as their first cause. However, currently, ideas like this or the search for specific characteristics of the creation process in general are only being refined by searching for the necessary indicators, for example, to determine the probability that the Universe was created consciously, and even come to the conclusion that such a probability is not zero (Broadbent, 2023), or they express skepticism about the possibility of clarifying relevant issues and reconciling ethical issues with the multiverse possibility (Pittard, 2023).

Again, Fredkin is only approaching the formation of those hypotheses that have become popular due to the active development of artificial intelligence. First of all, we are talking about the simulation hypothesis, which already in a more categorical manner asserts that the surrounding world is simulated by something more fundamental than anything we have encountered before, in particular by a superintelligence that we are unlikely to ever be able to comprehend. This simulation appears to be a purposeful design, while digital physics, as D. Chandler draws attention to, considers the emergence of the computational foundations of the Universe as a completely natural process that arose naturally (Chandler, 2023: 945). Hence,

the differences in explanations and possibilities of our cognition. Fredkin does not need the simulation hypothesis; it is enough for him to limit himself to the hypothesis of finite nature and to admit that some things will remain incomprehensible even when powerful artificial intelligence is created. Nick Bostrom, one of the supporters of the simulation hypothesis, on the contrary, argues that if the evolutionary development of humanity gradually takes place and subsequent civilizations will be able to achieve incomparably greater computing power than we currently have at our disposal, then it is quite reasonable to assume the existence of some kind of computer simulation in which we find ourselves (Bostrom, 2003: 255). So, our reality may not just have computational processes at its core, but even more – by its very nature, it may be interpreted as a simulation. If we accept that there is no difference at all between the virtual and the real, then, following D. Chalmers, we can conclude that virtual reality is the ultimate reality (Chalmers, 2022). And we can ask the question differently, as R. Yampolskiy does: if we accept that we live in a simulation, is it possible to escape from it, or rather go beyond its limits, and learn about what the real world is like (Yampolskiy, 2023).

But does the space of virtual reality have anything in common with the Other that Fredkin writes about? It is important for him to emphasize that the Other possesses an intellect of a fundamentally different nature and capacities than the intellect possessed by human beings. He writes:

"The point is that machines with Artificial Intelligence (AI) that we may build in the future will be qualitatively different in intellectual capability than are humans. We can't speak of AI as 'intelligent' in human terms. Something that may have created our universe in a purposeful fashion is very different than something with human-like intelligence or the kind of intelligence we will see in Earthbound AI. There might be no more in common than the concept of questions and the concept of finding answers" (Fredkin, 1993: 120).

In other words, Fredkin argues that artificial intelligence, in its principles of functioning, will resemble the intelligence that is generally inherent in biological species on Earth. The behavior of these species is rational and can be explained, and causal relationships can be traced in their actions, etc. Artificial intelligence is also, in principle, understandable to people in its choices and completely predictable, because it acts according to given algorithms. Artificial intelligence is also, in principle, understandable to people in its choices and completely predictable, because it acts according to certain algorithms. And even if it significantly surpasses a person in its capacities (in the course of scientific progress, any technical means each time surpass human abilities – whether physical or cognitive, related to memorization, performing computation, and now solving logical problems), this will not change anything in the matter of understanding the nature of the Other's intelligence, which is qualitatively different from everything we can deal with.

However, Fredkin's reasoning, especially in light of the latest technological advances in artificial intelligence, may be partially revised, since in many aspects, some things remain unclear to him. He writes about artificial intelligence as a process of computation, but says nothing clearly about the possibility of constructing artificial consciousness. It should become an integral part of the superintelligence, and it should already be asking what it means to be "created by the Other." A person may find themselves in the position of the Other in relation to the host of artificial consciousness. From his decisions and considerations obtained in the process of virtual reality simulation, it will be possible to understand a little more about the Other itself, which could be not only external to us, but also emergent in relation to the inner world of those who possess consciousness. Even if the quality of the Other, as Fredkin himself agrees, will remain incomprehensible to us, we can quantify the capabilities of the

Other – for this, we need to know the basic parameters necessary for computing the digital mechanics of our Universe and, accordingly, come closer to understanding its purpose, that is, what it was created for.

Of course, from the perspective of the digital process of creation, it was not so much the Universe that was created as the ultimate computer, which, like the ancient Greek arche, can be interpreted as the foundation of the world. Physics does not explain who set this algorithm, where and under what circumstances it was set, just as it does not explain anything that arose in the Other. In fact, from the standpoint of physics, we cannot say for what purpose this world was created, nor what place belongs to the human being within it. The only thing that can be logically assumed (but this statement would not contain any semantic information) is that: "In either case, our existence here on Earth might or might not be completely incidental to the purpose" (Fredkin, 1993: 120). However, Fredkin says a lot with this. Firstly, things like anthropocentrism or the anthropic principle are nothing more than hypotheses and should be questioned. It may well seem that our existence is not decisive in the project implemented in the Other. Secondly, the meanings embedded in the original intention may never be understood by us, and we will simply not be able to determine the nature of metaphysical truths, but only recognize that different answers to this question are possible (a multiplicity of interpretations). Asking the question about the first principles of the world may prove to be more important than answering it. Thirdly, no matter how uncertain a person's place in the world may be, and no matter how incomprehensible its foundations may be, what is important is that the world has a purpose. Fredkin is prompted to this idea by both the very idea of the finite nature and the possibility of regarding the foundation of our Universe as a vast digital computer that performs computation, as well as by the human capacity to construct computers. For any computer (constructed by humans or constituting the very foundation of the world) is universal by its nature, and therefore operates according to a predetermined algorithm in order to realize a final purpose that can be known only once the computation has been completed.

Knowledge of the inner world

At the same time, the question of whether it is worth looking for manifestations of the Other, and ultimately unique dimensions of being as such – within ourselves, inevitably makes us think about the prospects for digital analysis of the human inner world. Although Fredkin is not inclined to pose the question in this way, he conceives the digital approach as universal, and therefore it can be applied to the interpretation not only of the macrocosm, but also of the microcosm. However, the method that Fredkin himself chooses to analyze the microcosm, and accordingly to understand human beings as informational, seems unsatisfactory, since it concerns not so much the essence of things as the ways of using words. The meaning of the terms we use to describe a person's inner world, in his opinion, needs clarification. In other words, there is no need to invent new terms; we must understand what we are referring to with the terms we already have. Thus, we can reflect on the nature of the soul, as Fredkin himself does (Fredkin, 2000), we can ask whether computation is sufficient for thinking or the cognitive sphere, as Chalmers writes about (Chalmers, 2011), or, accordingly, for solving the mind-body problem, as Ralph Abraham and Sisir Roy do (Abraham & Roy, 2012), etc. Of course, there are many more terms that describe the inner world of a person: we are talking about the spirit, the microcosm, the mechanisms of self-monitoring, consciousness or self-awareness, and terms that have long become part of certain philosophical systems (such as nous, intellectus, Vernunft and Verstand) and which are accepted without evidence, without

conducting relevant experimental research. Fredkin proceeds from the assumption that what is important is not so much how to name internal states or processes, but rather their very essence and the extent to which it corresponds to what we mean by them. It is possible to correctly interpret the nature of things only by clearly realizing that any entities are preceded by a certain computational process.

Thus, from the standpoint of the information-computational approach, Fredkin writes that: "The soul is an informational entity, which is constructed out of the states and the arrangements of material things" (Fredkin, 2000: 5). That is, he does not accept the position of eliminative materialism, according to which terms that describe the inner world of a person but are not neuroscientific do not allow for adequate explanation. From his proposed definition of the soul, it becomes clear that it is a separate entity, which, like everything else in the world, has an informational nature. The soul becomes that structure of information which describes the cognitive states of memory, consciousness, and the way we perceive and understand the world. Its existence is shaped by the presence of a material body, but, in essence, it is an informational, not a physical, process and therefore an informational structure in a computing environment.

Fredkin does not reduce the soul to material processes: "…the soul is not made up of matter; it is an informational construct wherein the states and arrangements of matter and energy are used to represent the information" (Fredkin, 2000: 19). In other words, the nature of the inner world of a person (soul) can be explained as a pattern of different information states, as a kind of program that operates on the basis of a material substrate. Under such conditions, a person's inner world (in its conscious and unconscious dimensions) is a kind of virtual space, and information states as patterns can be reproduced on other hosts and within a certain virtual world. Moreover, these virtual worlds, as evidenced by the development of technologies under the conditions of creation by artificial intelligence, can become something different from what has been created by humans (Synytsia, 2025).

By proposing an information-computational approach to explaining the nature of the world, the cognitive sphere is reduced to the sphere of virtual reality, part of which, according to the simulation hypothesis, may be our consciousness. The distinction between mind and body, virtual and real, may well be reinterpreted within the information paradigm (for the world we think or imagine, in a certain sense, resembles a virtual world, and vice versa). In this case, vision is decisive. Thinking and extended substances become, respectively, aspects of a single informational entity (which is, in fact, analogous to the given in Bertrand Russell's concept of neutral monism). Under the conditions of information-computing processes as the fundamental basis of the world, the virtual no longer becomes one of the manifestations of reality, but, on the contrary, reality becomes a manifestation of the virtual. Within this virtual space, the soul (mental) acts as a stable information configuration that organizes data in such a way that consciousness (an information structure embodied at the local level) and, accordingly, self-awareness (a meta-information function, which is essentially information about information, an algorithm that explores itself) emerge.

The soul can organize data that comes from outside or is the result of introspective analysis because it is not homogeneous itself. This conclusion inevitably follows from the essence of his approach. After all, Fredkin himself argues that: "In order to understand our definition of '*souls'* we need to understand that a *soul* is made up of many parts, that not all *souls* in human bodies have all of the parts that other *souls* have and that parts of a human's *soul* can exist outside of that human's body…" (Fredkin, 2000: 4). He calls the parts of the soul *soullecules* – they are actually particles of information understood by a person (certain

information about something). We can share them with others both in the form of short units of meaning expressed through words, in musical compositions, etc. The learning process enriches our soul with new particles, but sometimes the old ones are lost. In this way, the soul changes every time. The combinations of mental states and processes each time represent the uniqueness of each person's inner world.

To explain how the cognitive sphere functions, Fredkin distinguishes between a dynamic and a static soul. About the first one he says: "A dynamic soul is just what we normally think of as the soul of a living, awake and active being" (Fredkin, 2000: 5), and about the second: "A static *soul* is simply digital information that is similar to a combination of computer programs and computer data" (Fredkin, 2000: 5). As we can see, he actually identifies the dynamic soul with the conscious processes that we experience when we contemplate the world, reflect on it, feel, learn, communicate, etc., and the second (static soul) with the set of knowledge that is stored in our memory and can be activated and used when needed as a component of the dynamic soul. When a person, for example, thinks about proving the Pythagorean theorem, listens to a musical composition by Mozart, or empathizes with one of the literary characters invented by Shakespeare, he actually begins to be in the same state as others when they think about it; all of them actualize in themselves that part of the soul that the corresponding authors once experienced and expressed in their works.

Essentially, a dynamic soul is an activated program that performs computation and uses the body to interact with the environment and make itself known to others. Then the static soul actually corresponds to a set of digital data that is stored on separate media and can be activated each time the program is run. It can only be run in a specific environment. In this case, it is not at all about the uniqueness of this environment, but rather that it satisfies the appropriate conditions: just as a program can be run on different computers, so the soul (its individual parts) can be saved and reproduced. This requires not so much a human body as an environment that will perform all the biological functions of the organism that are necessary for the functioning of the soul. That is, the environment is compatible with the soul, in Fredkin's interpretation, it is some kind of perfect computer. He explicitly states this: "We believe that a contemporary high end personal computer can have the computational resources to host the dynamic soul of one person in real time" (Fredkin, 2000: 21). However, he immediately clarifies that in order to understand the soul, one must first properly understand how the body functions and what the environment in which it functions should be. There will always be a question of how to interact with the soul – and in this matter, Fredkin draws an analogy with the study of the principles of operation of a RAM chip, the study of which, from the outside, tells us nothing about what it is capable of, but there are many methods for discovering what it does.

Contemporary research suggests not only assessing the hardware resources necessary for this kind of manipulation of the soul, but also developing specific operational indicators of consciousness in artificial systems that are not achievable in modern systems but are technologically achievable in general (Butlin et al., 2023). Although learning systems can acquire "multiple selves," evolutionarily, this makes artificial systems more flexible and adaptable to changing environmental conditions (Dulberg et al., 2023). At the same time, such metaphysical issues related to the multiple copying and transfer of personality (the problem of multiplicity), as some researchers argue, can be completely solved not only at the level of an artificial personality, but also at the level of a human personality (as a set of manifested qualities of the soul), which can also develop in different ways (Weber, 2025).

For his part, Fredkin, taking into account the imperfection of artificial intelligence

technologies at the time, sought to focus not so much on the technical aspects of this issue as on the conceptual and philosophical ones, and therefore, in his characteristic manner (giving new meaning to concepts), he tries to reconcile basic tenets of digital cosmology with the key tenets of religious and worldview teachings. For example, he says that: "Reincarnation is possible, but not in ways previously imagined; the soul may, in partial or nearly total form, survive the death of the body; the partial or nearly total soul may be able to exist separate from the body; the soul is potentially an immortal informational construct; thus a part of physics" (Fredkin, 2000: 4). He is convinced that "the laws of physics do not prohibit the survival of a person's soul, largely intact, after the death of that person's body", but immediately adds: "…we do not believe that such an event has yet occurred" (Fredkin, 2000: 4). The possibility of preserving the soul immediately raises the question of the immortality of the soul, which again seems possible to Fredkin, but this cannot be done in violation of the laws of physics: "For a soul to be immortal, it must either be resident in an immortal host, or it must be able to survive the mortality of each member of a sequence of hosts. Since the soul is an informational construct" (Fredkin, 2000: 26).

For this possibility to become a reality, the development of digital technology must obviously take place. Even today, decades after Fredkin first articulated his reflections on digital philosophy, we can only talk about positive dynamics, not about any specific results; about understanding how the relevant processes might occur, not about how to implement them. For example, from the digital perspective, reincarnation is interpreted as the possibility of expressing the same mental states through a host of a different nature. It can be assumed that if it is possible to transfer states of consciousness into a virtual reality, then the reverse process is not excluded – when these states are reproduced on other hosts, not only digital, but also biological (in this case, we will be dealing not with escape in Yampolskiy's interpretation above, but with return). Fredkin's approach does not exclude such a possibility; rather, it regards its occurrence as inevitable.

However, it is precisely this transition that, in many respects, appears to be an insoluble difficulty in contemporary cognitive theory and the main point of criticism of reductionism. Qualia, or the subjective characteristics of sensory experience, as anti-reductionists argue, cannot be reproduced: they are unique in nature. Attempts to make the real part of the virtual rest on analogy; reality is much more complex. So the digital approach, if it ever claims the status of a theory of everything and explains the nature of fundamental interactions, must also explain the cognitive sphere. Fredkin's attempt seems noteworthy, and with the development of artificial intelligence, interest in the information-computational approach has only increased, while human existence, expressed in the correlation of mind and body, continues to remain a mystery, to unravel which, the following more substantiated explanations are needed, not only theoretical, but also at the empirical level. Their attainment would be made possible, in particular, by artificial intelligence technologies, the development of which will demonstrate how reasonable the ideas of Fredkin's digital philosophy are and how they can be used to explain the nature of reality and the human being.

🕮 References

Abraham, R., and Roy S. (2012) A digital solution to the mind/body problem. In *P. Simeonov, L. Smith, A. Ehresmann (Eds.), Integral Biomathics* (pp. 213-225). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-28111-2_20

Bostrom, N. (2003) Are we living in a computer simulation? *Philosophical Quarterly,* 53(211), 243-255. https://doi.org/10.1111/1467-9213.00309

Broadbent, D. (2023) Could our Universe have been initiated intentionally: An epistemological framework. *Philosophy and Cosmology*, 31, 4-12. https://doi.org/10.29202/philcosm/31/1

Butlin, P., Long, R., Elmoznino, E., Bengio, Y., Birch, J., VanRullen, R. et al. (2023) *Consciousness in artificial intelligence: Insights from the science of consciousness*. arXiv. https://arxiv.org/abs/2308.08708

Chalmers, D. (2011) A computational foundation for the study of cognition. *Journal of Cognitive Science,* 12(4), 325-359. https://doi.org/10.17791/jcs.2011.12.4.325

Chalmers, D. (2022) *Reality+ : Virtual worlds and the problems of philosophy*. New York: W. W. Norton & Company.

Chandler, D. L. (2023) Could the Universe be a giant quantum computer? *Nature,* 620 (7976), 943-945. https://doi.org/10.1038/d41586-023-02646-x

Dulberg, Z., Dubey, R., Berwian, I. M., & Cohen, J. D. (2023) Having multiple selves helps learning agents explore and adapt in complex changing worlds. *Proceedings of the National Academy of Sciences*, 120(28), e2221180120. https://doi.org/10.1073/pnas.2221180120

Faizal, M., Krauss, L., Shabir, A., & Marino, F. (2025). Consequences of undecidability in physics on the theory of everything. *Journal of Holography Applications in Physics*, 5(2), 10-21. 10.22128/jhap.2025.1024.1118

Fredkin, E. (1990) Digital mechanics: An informational process based on reversible universal cellular automata. *Physica D: Nonlinear Phenomena,* 45(1-3), 254-270. https://doi.org/10.1016/0167-2789(90)90186-S

Fredkin, E. (1992) Finite nature. *Proceedings of the XXVIIth Rencontre de Moriond* (Les Arcs, Savoie, France, March 22–28, 1992) (pp. 345-354). Gif-surYvette, France: Editions Frontieres.

Fredkin, E. (1993) A new cosmogony. In D. Matzke (Ed.), *Proceedings of the Workshop on Physics and Computation* (*October 2–4, 1992)* (pp. 116-121). New York: Institute of Electrical and Electronics Engineers Computer Society Press. https://doi.org/10.1109/PHYCMP.1992.615507

Fredkin, E. (2000) *On the soul* (Unpublished manuscript). https://www.scribd.com/document/235105454/On-the-soul-Ed-Fredkin

Fredkin, E. (2003) An introduction to digital philosophy. *International Journal of Theoretical Physics*, 42, 189-247. https://doi.org/10.1023/A:1024443232206.

Fredkin, E. (2004) Five big questions with pretty simple answers. *IBM Journal of Research and Development*, 48(1), 31-45. https://doi.org/10.1147/rd.481.0031

Fredkin, E. (2005) A computing architecture for physics. *CF'05: Proceedings of the 2nd conference on Computing frontiers,* 273-279. https://doi.org/10.1145/1062261.1062307

Pietarinen, A. V. (2025). Synechism 2.0: Contours of a new theory of continuity in nature. *Studies in History and Philosophy of Science Part A,* 100, 103–118. https://doi.org/10.1016/j.shpsa.2025.05.006

Pittard, J. (2023) The Skeptical challenge of the theistic multiverse. *Ergo*, 9, 36. https://doi.org/10.3998/ergo.2925

Steinhart, E. (1998) Digital metaphysics. In T. W. Bynum & J. H. Moor (Eds.), *The Digital Phoenix: How Computers are Changing Philosophy* (pp. 117-134). Cambridge: Blackwell.

Synytsia, A. (2025) Remarks on Karl Popper's idea of three worlds in the context of the development of virtual reality technologies. *Actual Problems of Mind*, 26, 31-48. https://

doi.org/10.55056/apm.7748

Weber, C. (2025) The multiplicity objection against uploading optimism. *Synthese* 205, 226. https://doi.org/10.1007/s11229-025-05057-9

Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus* / trans. by C. K. Ogden. London, Kegan Paul, Trench, Trubner & Co.

Yampolskiy, R. (2023) How to escape from the simulation. *Seeds of Science*. https://doi.org/10.53975/wg1s-9j16